Parameter Estimation using Median Ranked Set Sampling

by

Abdul-Baasit Shaibu

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

MATHEMATICAL SCIENCES

December, 1999

Parameter Estimation using Median Ranked Set Sampling

BY

Abdul-Baasit Shaibu

A Thesis Presented to the
FACULTY OF THE COLLEGE OF GRADUATE STUDIES
KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE
In
Mathematical Sciences

December, 1999

UMI Number: 1401870

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS
Dhahran 31261, SAUDI ARABIA

COLLEGE OF GRADUATE STUDIES

This thesis, written by Abdul-Baasit Shaibu under the supervision of his thesis advisor and approved by his Thesis committee, has been presented to and accepted by the Dean of the College of Graduate Studies, in partial fulfillment of the degree of MASTER OF SCIENCE IN MATHEMATICS.

Thesis Committee

Dr. Hassen Muttlak
Chairman

Dr. Walid S. Al-Sabah
Member

Dr. Anwarul Haque Joarder
Member

Dr. Mohammad Iqbal
Member

Dr. Walid S. Al-Sabah
Department Chairman

Dr. Abdullah M. Al-Shehri
Dean, College of Graduate Studies

Date: 18-12-99

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

To

My Parents and *Auntie Grace*

Acknowledgement

Thanks and praise are due to God Almighty, the Sustainer of the worlds, with whom all things are possible.

I also wish to thank King Fahd University of Petroleum and Minerals (KFUPM) for giving me the opportunity to study for the Master of Science degree. I am also thankful to Dr. Hassen Muttlak, my thesis advisor, for selflessly and patiently sharing his enormous research experience with me. To Dr. Walid S. Al-Sabah, Dr. A. H. Joarder and Dr. M. Iqbal, all of my thesis committee, I extend lots of gratitude for the encouragement they gave me in the course of my research. I am extremely grateful to Dr. Walid, who as the chairman of the Department of Mathematical Sciences, did a lot to ensure my ease and well being. I acknowledge my indebtedness to Mr. Yushau Balarabe, who took me through the formalities that would have been too difficult for me to go through on my own when I first arrived at KFUPM. I thank him very much for holding me close to his heart.

I extend my absolute gratitude to Dr. Robert Huotari, a former member of faculty I was supposed to have worked with for taking me as a buddy, for completely identifying with me and for teaching me things too numerous to mention here. He among other things encouraged me to learn computer programming, which has been my most powerful tool in this work. I greatly acknowledge Mr. Taofeek Ayinde, a graduate student of Mechanical Engineering and a sincere friend for teaching me for free, how to write computer programs. I am also very grateful to Dr. Gary R. Miller who after the departure of Dr. Huotari, did not rest until he was sure that I had found something I was happy to do for my thesis. I also thank him for wonderfully acquainting with me. I greatly appreciate the effort of Professor M. A. Rahim of University of New Brunswick, Canada in improving my research capability when he taught me Advanced Quality Control as a Visiting Professor to KFUPM. I sincerely thank Professor Stokes for promptly sending me one of her publications when I asked for it.

I thank all the members of Faculty of the Department of Mathematical Sciences, particularly those who taught me courses for credits or otherwise. I must also thank Professor D. A. Akyeampong, Dr. S. A. Adjei, Dr. N. K. Kofi Nti and Dr. D. Adu-Gyamfi all of the Department of Mathematics, University of Ghana, who under very difficult conditions, gave of their best to prepare me well enough to qualify for graduate studies. A similar acknowledgement is due to Dr. S. I. K. Odoom, Dr. F. K. Atsem, Dr. I. G. Ackar and Mrs. Danso Manu all of the University of Ghana, and Mr. H. Q. B. Donkor now of SOS-Herman Gmeiner International college who were responsible for my undergraduate education in Statistics. I must also acknowledge Mr. Saadat, Mr. Z. Qamar, Mr. Shareef and the rest of the members of staff of the Department of Mathematical Sciences for their cooperation and understanding.

Finally, I am grateful to the entire Ghanaian community in KFUPM and Saudi ARAMCO, and the friends of the entire African community for their prayers, encouragement and support in various ways. Very special to me are my parents, Auntie Grace, Mr. Kwabena Akyeampong and Uncle B. K. Gado.

Abdul-Baasit Shaibu.
December 1999.

Abstract

Name: Abdul-Baasit Shaibu
Title: Parameter Estimation using Median Ranked Set Sampling
Major Field: Mathematics
Date: December, 1999

Ranked set sampling was introduced and used as a cheap and efficient method of estimating mean pasture yield in 1952. Since the efficiency of this method is susceptible to errors in ranking, several modifications, including median ranked set sampling and extreme ranked set sampling, have been proposed and have been shown to improve the efficiency of estimation in some cases. Ranked set sampling is parametric in nature. However, several authors have used it in estimating the parameters of various distributions and have shown their estimators to be more efficient when compared to those based on simple random samples of the same size.

In this thesis, we use the methods of maximum likelihood estimation and linear unbiased estimation based on median ranked set sampling and extreme ranked set sampling to estimate the parameters of several distributions. We show that in most cases, estimators based on median ranked set samples are more efficient than those based on ranked set samples and simple random samples of the same size. In estimating the normal standard deviation, we show that the maximum likelihood estimator based on extreme ranked set samples does better than those based on ranked set samples and median ranked set samples of the same size.

MASTER OF SCIENCE DEGREE

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS
DHAHRAN, SAUDI ARABIA
DECEMBER, 1999

خلاصة الرسالة

اسم الطالب الكامل: عبد الباسط شيبو

عنوان الرسالة : تقدير المعلمات باستخدام معاينة المجموعة المرتبة وسطياً

التخصص : الرياضيات

تاريخ الشهادة: رمضان 1420

لقد اقترحت و استخدمت طريقة معاينة المجموعة المرتبة في عام 1952م كطريقة رخيصة و فعالة في تقدير معدل محصول المراعي. بما إن فعالية هذه الطريقة معرضة للنقصان بسبب الخطأ في ترتيب الوحدات فلقد تم اقتراح تعديلات كثيرة من بينها معاينة المجموعة المرتبة وسطياً و معاينة المجموعة المرتبة تطرفياً. وأن هذه التعديلات برهن على فعاليتها في تقدير الو ط الحساب للمجتمع في بعض الحالات. إن طريقة معاينة المجموعة المرتبة هي طريقة غير معلميه بطبيعتها و لكن كثيراً من الكتاب استخدمها لتقدير المعلمات لكثير من التوزيعات وأثبتوا أن تقديراتهم اكثر فعاليةً مقارنة بالتقديرات المعتمدة على المعاينة العشوائية البسيطة وباستخدام نفس حجم العينة.

في هذه الرسالة سوف نستخدم طريقة (M.L.E.) و (Blue) لتقدير معلمات بعض التوزيعات باستخدام معاينة المجموعة المرتبة وسطياً و معاينة المجموعة المرتبة تطرفياً. لقد برهنا في اغلب الحالات إن التقديرات المعتمدة على معاينة المجموعة المرتبة وسطياً إنها اكثر فعاليةً من التقديرات المعتمدة على معاينة المجموعة المرتبة و المعاينة العشوائية البسيطة وباستخدام نفس حجم العينة. لقد برهنا إن تقدير الانحراف المعياري للتوزيع الطبيعي باستخدام طريقة (M.L.E.) اكثر فعاليةً باستخدام معاينة المجموعة المرتبة تطرفياً من التقديرات المعتمدة على معاينة المجموعة المرتبة أو معاينة المجموعة المرتبة وسطياً و بنفس حجم العينة.

درجة الماجستير في العلوم الرياضية

جامعة الملك فهد للبترول و المعادن

الظهران، المملكة العربية السعودية

رمضان 1420 /ديسمبر 1999

Contents

List of Tables

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1. Introduction

In most sample surveys, a reasonable number of the sampling units can be fairly accurately ordered with respect to a variable of interest without actual measurement and at little or no cost. On the other hand, the direct measurement of these units may be very tedious and/or expensive. Examples of such situations include the estimation of the mean pasture or forage yield, the estimation of the pollution level of the environment, the estimation of the mean height of trees, etc. Considering the estimation of mean pasture yield, McIntyre [19] proposed the method of ranked set sampling (RSS) as an alternative to the method of simple random sampling (SRS). The RSS method consists of first drawing n random samples from the population, each of size n, and ranking each of the n samples by visual or judgement process. Then the smallest observation from the first sample is chosen for measurement, as is the second smallest observation from the second sample. The process continues in this way until the largest observation from the n^{th} sample is measured, yielding a total of n measured units, one from each ordered sample. The entire cycle is repeated m times until a total of mn^2 units have been drawn from the population and mn have been measured. The mn measured observations form the 'ranked set sample'. We will explain this method of sampling in detail in the next Chapter along with other sampling methods. As an estimator of the population mean μ, McIntyre [19] asserted without supporting mathematical theory that the sample mean from RSS is more efficient than the sample mean from SRS. That is to say, the RSS mean has smaller variance than the SRS mean. Thereafter, a lot of work has been done in this area, some of which we review below.

In this thesis, we investigate the maximum likelihood estimators of the location-scale family of distributions under MRSS as has been done by Stokes [45] under RSS. We also propose some unbiased estimators of the same parameters and investigate their performance.

In the next chapter, we outline the basic concepts required for the proper understanding of work done in this thesis. We find the maximum likelihood estimates of the location and scale parameters under MRSS in Chapter 3 and investigate their performance by finding the Fisher information for the parameters under consideration and using the lower bound variance to compare the performance of the estimators under MRSS and those under SRS. Chapter 4 is devoted to some unbiased estimators of the location-scale parameters under MRSS. We propose these estimators by modifying the best linear unbiased estimators in Stokes [45]. In Chapter 5, we consider maximum likelihood estimation under extreme ranked set sampling (ERSS) following the methods of Chapter 3. Finally in the last chapter, we summarize and discuss the results of the entire thesis.

1.2. Literature Review

Since 1952 when McIntyre intuitively proposed and used the method of ranked set sampling, it hasn't been used until 1966 when Halls and Dell [11] applied it in the estimation of forage yield. The method went into recess again until 1968 when Takahasi and Wakimoto [50], independent of the previous work, described it and supplied a rigorous mathematical theory, which supports McIntyre's earlier intuitive assertion. Thereupon, there has been a lot of work in the area with very interesting modifications. We classify and review some of these works under non-parametric methods, parametric methods, hypotheses testing and regression analysis.

The RSS procedure is referred to as the MTW procedure in recognition of McIntyre [19] and Takahasi and Wakimoto[50]. We will sometimes refer to the estimators resulting from this method as the MTW estimators.

1.2.1. Non-Parametric Methods

Takahasi [47] considered the problem in Takahasi and Wakimoto [50] under the situation where the elements within each set are correlated. He presented a model and an estimator of the population mean, and gave the efficiencies of the estimators for some specific distributions. Takahasi [48] followed up with a modification of RSS where an element is randomly selected and measured before the determination of its rank. Thus, the unit to be quantified is not predetermined by its position.

Dell and Clutter [10] considered the case of erroneous ranking with respect to a variable of interest and showed that the efficiency of the RSS statistics decreases with increasing ranking errors. They however proved that even with the errors in ranking, the RSS is still more efficient than the SRS and that in cases where judgement is so poor as to produce a random sample, then RSS is as efficient as SRS.

Yanagawa and Shirahata [52] constructed an estimator of the population mean based on elements selected subject to a selective probability matrix (SPM); an m by n matrix of probabilities $\{P_{ij}: i=1, 2, \ldots, m; j=1, 2, \ldots, n\}$ satisfying $\sum_{j=1}^{n} P_{ij} = 1$ for each $i = 1, 2, \ldots, m$. The procedure consists of selecting at random, m sets each of n elements. The elements within each set are ordered and the i^{th} order statistic of the j^{th} set is selected for quantification with probability P_{ij} from the SPM. The mean of the selected units is called the YS estimator of the population mean. They proved that this estimator is unbiased if the SPM satisfies $\sum_{i=1}^{m} P_{ij} = m/n \quad j = 1, 2, \ldots, n$. The estimator is also shown to be a generalization of the MTW estimator. The dominance of the estimator over the SRS estimator was proved.

Shirahata [39] proposed the concept of selective probability, which generalizes the YS procedure. Based on this, he found uniformly unbiased procedures for some (m, n), where m and n respectively denote the number of row and columns of the probability matrix. He further considered symmetric distributions that do not allow uniformly unbiased procedures and found unbiased procedures for them for some (m, n).

Stokes [41] considered a situation where a variable of interest, say X, cannot easily be ordered but a characteristic Y correlated with X can. She proposed a sampling method based on the ranks of the Y's (the concomitant variables). Bivariate samples (X, Y) are selected following the MTW method. The Y's are ranked and the X associated with the smallest Y in the first sample is quantified as is the X associated with the second smallest Y in the second sample. This continues until the X associated with the largest Y in the last sample is measured. The precision of the mean of the quantified X's as an estimator of the population mean depends on the strength of the correlation between the X's and the Y's. If the correlation coefficient is one, then the estimator of X is equivalent to the MTW estimator. On the other hand, if the X's and Y's are completely independent, we get the SRS estimator.

Stokes [42] studied the estimation of the population variance, σ^2 using RSS with n sets and m cycles. She proposed an estimator of σ^2, which is generally biased. The estimator was however shown to be approximately unbiased for large values of either the cycle size, m or the set size, n or both. The impracticability of large values of n was however pointed out and the enlargement of sample size was restricted to large values of m. The relative precision $RP = \frac{Var(s^2)}{MSE(\hat{\sigma}^2)}$ is shown to be more than or equal to one as $mn \rightarrow \infty$, where s^2 and $\hat{\sigma}^2$ respectively represent the SRS and RSS estimators of σ^2. Equality holds in cases where judgement ranking is poor enough to produce a simple random sample.

Yanagawa and Chen [51] studied the MG-procedure; an alternative to the YS-procedure described by Miller and Griffiths (unpublished). In this method, $n = 2r$ sets each of size m (not necessarily equal to n) are selected and each set ranked by judgement. A set of r independent random variables $I_1, \ldots, I_r$ are considered such that I_j is selected with probability P_{ij} *for* $i = 1, \ldots, r; j = 1, \ldots, m$ with the P_{ij}'s satisfying the conditions of the YS-procedure. For each $i = 1, \ldots, r$, if $I_i = i_j$, then the pair consisting of the i_j^{th} order statistic in the j^{th} set and the $(m+1-i_j)^{th}$ order statistic of the $(n+1-j)^{th}$ set, $\{X_{j(i_j)}, X_{n+1-j,(m+1-i_j)}\}$, is selected and quantified. This yields a sample of size n, the mean of which is the proposed estimator of the population mean. This estimator is shown to be unbiased if for all j, $\sum_{i=1}^{r}(P_{ij} + P_{i,(m+1-j)}) = m/n$. It was shown that the MG procedure is more efficient than the YS and MTW procedures when $m \neq n$, and that all the three procedures agree when $m=n$.

In Ridout and Cobby [35], attention was given to ranking errors and the fact that the need to rank the elements in a set may render random sampling inoperative. The effect of these on the precision of the RSS estimator was examined. An example was given to show that the reduction in the relative precision is due more to the non-random sampling than to errors in ranking.

Stokes and Sager [46] characterized the RSS method and applied their characterization to the estimation of the cumulative distribution function (cdf) of a population. The RSS empirical distribution function is shown to be unbiased for the population cdf and also more efficient than that of the SRS even in the presence of ranking errors. They also discussed how to use the Kolmogorov-Smirnov statistic and a RSS empirical distribution function to improve the confidence bands for the cdf over the SRS bands.

Muttlak and McDonald [29] proposed a two-phase sampling procedure with a naturally size-biased first phase. That is, the probability of selection in the first phase is proportional to size for each unit. They employed RSS to obtain second phase sample and showed that this increased the efficiency of the estimators of the population size and mean. They also considered the effect of erroneous ranking on the efficiency of the estimators and showed that in spite of the reduction in the efficiency, their proposed procedure is still dominant over simple random sampling. Muttlak [26] considered this problem using median ranked set sampling (MRSS); a sampling plan that would be explained in detail in the next chapter.

Kvam and Samaniego [16] studied the non-parametric maximum likelihood estimator of the distribution function and proved its existence and uniqueness. They presented a general numerical procedure and showed that it converges to their proposed estimator. They modified their procedure to be applicable when the RSS is unbalanced and the procedure proposed by Stokes and Sager [46] is not applicable. They further showed that if the ranked set sample is balanced, their procedure does better than that of Stokes and Sager [46].

Muttlak and McDonald [30] used RSS with size biased probability in combination with the line intercept method to improve the efficiency of the usual unbiased estimators of cover, density and total amounts of variables under study. They proposed a two-stage procedure. In the first stage, the line intercept method was used to select units whilst in the second stage, the method of RSS was combined with size biased probability to construct the estimators. They proved the unbiasedness of their proposed estimators and illustrated their dominance over the usual ones with a practical example. Muttlak [21] also studied a similar subject and applied the procedure in the estimation of coverage, density and total number of stems per unit area of rose rock (*Cistus villosu*) in a study area in Jordan.

Kvam and Samaniego [15] showed that the empirical averages, as estimators in RSS are inadmissible. They identified several situations in which the performance of the RSS estimators can be uniformly improved. They also provided sufficient conditions for the inadmissibility and proposed estimators that extend to the unbalanced case of RSS.

Patil et al [32] studied RSS from a finite population which follows a linear or quadratic trend. They found expressions for the variance and relative precision of the RSS estimator for several set sizes and showed that the dominance of RSS over SRS increases with decreasing population size. In comparing RSS to stratified and systematic sampling, they showed that it is dominant in some cases.

Bohn [2] studied the various non-parametric settings of RSS namely; the RSS empirical distribution function, the two-sample location setting, the sign test and the signed-rank test and discussed the similarities and differences of the RSS procedures for each of these settings.

In Muttlak [23], a modification of the RSS method, pair ranked set sampling was introduced. In this method, he defined $k = n/2$ for even set size, n, and $L = (n+1)/2$ for odd set size. L or k independent samples, each of size n are taken according as n is odd or even. Ranking the units in each of the samples with respect to the variable of interest, the i^{th} smallest and the i^{th} largest unit in the i^{th} sample are selected for quantification. In case of odd n, the L^{th} largest unit and the L^{th} smallest unit coincide with the L^{th} order statistic (the median), which is chosen for measurement. In the case of even n, the k^{th} smallest unit and the k^{th} largest unit coincide with the k^{th} and $(k+1)^{st}$ order statistics respectively. Each case results in a quantified sample of size n. As in the case of RSS, the cycle may be repeated m times to yield a quantified sample of size nm. Estimators of the population mean under this sampling plan were proposed and shown to be dominant over the corresponding SRS estimator.

Muttlak [25] proposed median ranked set sampling (MRSS) as a modification of the ranked set sampling (RSS) to reduce loss of efficiency in RSS due to errors in ranking and as an improvement upon the efficiency of the estimator of the population mean. He proposed an estimator of the population mean. which is unbiased for symmetric distributions and biased otherwise. He clearly demonstrated with several distributions that in all cases, his estimator does better than the MTW estimator. He also studied the effect of ranking errors in reducing the efficiency of the estimators under the MRSS method.

Patil et al. [33] categorized the work done in the area of ranked set sampling into three: theory, methods and applications. They reviewed these aspects under the same notation and illustrated the performance of RSS against SRS in the determination of the level of contamination at a hazardous waste site. They also demonstrated the use of RSS methods in improving the formation of composite samples.

1.2.2. Parametric Methods

Sinha et al. [40] used RSS and some modifications of it to estimate the parameters of the normal and exponential distributions. Unlike previous considerations, which are non-parametric, their article assumed partial knowledge of the underlying distribution without any knowledge of the parameters. For each parameter, they proposed best linear unbiased estimators (BLUEs) for the full and partial RSS. In the case of the partial RSS, they found the least number of cycles for which the proposed estimators dominate the SRS estimators. Lam et al. [17] considered the two-parameter exponential family in the same spirit as in Sinha et al. [40]. They focused on all the parameters of the family including the quantiles.

Another work on parametric ranked set sampling is by Stokes [45] who considered the location-scale distribution, $F\left(\frac{x-\mu}{\sigma}\right)$, and estimated μ and σ using the methods of maximum likelihood estimation (MLE) and best linear unbiased estimation. A general method for finding BLUEs of these parameters using RSS was studied. For some distributions, the BLUEs were found to be nearly as efficient as the MLEs, while for others the BLUEs did badly.

Bhoj and Ahsanullah [6] estimated the parameters of the generalized geometric distribution and showed that their estimators are more efficient than those derived using ordered least squares.

Bohj [5] obtained linear unbiased estimators of the location and scale parameters of the extreme value distribution using RSS. He showed that these estimators are more efficient than the ordered least square estimators. He also showed that his estimator of the population mean is dominant over that of the usual RSS estimator.

Patil et al. [34] studied among other things, the performance of RSS and showed that it is a monotone increasing function of set size for the wide class of ranking models that satisfy coherence; ranking on a set is consistent with ranking on every superset.

Takahasi and Futasuya [49] considered samples from a finite population without replacement and studied the concepts of likelihood ratio dependence and negative regression dependence. They demonstrated the superiority of RSS estimators to those of SRS.

Hossain and Muttlak [13] showed that the estimators of the population mean and standard deviation under paired ranked set sampling are more efficient than those under SRS, RSS, and even the minimum variance linear unbiased estimators (MVLUE). Restricting their study to the normal distribution with erroneous ranking, paired ranked set sampling was still shown to be dominant over the rest of the schemes.

1.2.3. Tests of Hypotheses

Bohn and Wolfe [3] used the empirical distribution function for ranked set samples to develop non-parametric inference techniques for RSS data in the two-sample location problem. Using the ranked set empirical distribution function, they proposed estimation and testing procedures that do not depend on any particular distribution. They showed that this approach could lead to a more powerful version of the Mann-Whitney-Wilcoxon (MWW) procedure than the SRS version. Bohn and Wolfe [4] followed up with the examination of the effect of imperfect judgement rankings on the properties of the MWW procedures.

Based on the estimators of the normal mean, μ, derived in Sinha et al. [40], Shen [38] used the concept of RSS and derived tests for μ when the variance is known. He showed that his proposed tests are more powerful than the traditional normal test.

Abu-Dayyeh and Muttlak [1] showed that the tests based on RSS are more powerful than the uniformly most powerful test (UMPT) and the likelihood ratio test (LRT) on the scale parameter of the exponential distribution under SRS. A similar conclusion was drawn in the case of the UMPT on the scale parameter of the uniform distribution.

Muttlak and Abu-Dayyeh [28] showed that the tests for the normal mean and variance under RSS are more powerful than those under SRS.

Koti and Babu [14] studied the sign test under RSS and showed for some continuous distributions that this test is more powerful than a similar test under SRS. They also discussed the effect of imperfect judgement on the test and concluded that it may result in a higher type I error probability for the RSS sign test than for the SRS sign test.

1.2.4. Regression Analysis

Patil et al. [31] compared the RSS estimator and the regression estimator on the basis of the correlation between a variable of interest and its concomitant variable assuming that both variables follow a bivariate normal distribution. They showed that the regression estimator is more efficient if the correlation is high (greater than 0.85) while the RSS estimator is more efficient when the correlation is low (less than 0.85). Yu and Lam [57] assumed that the population mean of the concomitant variable is known. Otherwise, they proposed double sampling to estimate the population mean. They also considered the effect of the shape (or symmetry) of the distribution of the concomitant variable on the performance of their approach.

Muttlak [20] studied parameter estimation in simple linear regression under RSS. He showed that ranking either on the dependent or on the independent variables facilitated the dominance of the RSS estimators over their SRS counterparts. It was however shown that when ranking is on the independent variable and the correlation between the dependent and independent variables is low (less than 0.25), then the RSS procedures are not useful.

Considering the one-way layout model, Muttlak [22] used RSS to improve upon the efficiency of parameter estimators under SRS. Muttlak [24] showed that under RSS, estimators of the parameters of a multiple regression model are dominant over those under SRS. Using MRSS and basing the ranking on a concomitant variable, Muttlak [27] estimated the population mean of a variable of interest. He showed that this approach is more efficient as compared to the method of RSS. He also showed that his approach dominates the approach of regression estimators on condition that the correlation between the auxiliary variable and the variable of interest in the regression model is above 0.9. Samawi and Muttlak [36] demonstrated the results of Muttlak [24] in the case of the ratio estimator.

CHAPTER 2

SAMPLING METHODS AND PRELIMINARIES

In all investigations, there is usually a set or a collection that investigators wish to draw conclusions about. Examples of such sets include the total number of human beings in a well-defined geographical area, the output of industrial machines, agricultural yield, etc. Any such set is referred to as a target population or simply, a population. It is clear from these examples that the size of a population is usually enormous. Thus, studying every single unit of a given population may be too costly, tedious and time consuming, or impossible. Hence, investigators usually have ways of properly selecting and studying only a portion of their target population, so that their conclusions are applicable to the entire population. The process of selecting only a portion of a population for study is known as random sampling, the selected portion being a random sample. The use of the characteristics of a random sample to represent those of a population is known as estimation, the random sample characteristics being estimates of the population characteristics. For example, a random sample of one thousand people in a country of twenty million may be investigated in an opinion survey on a given policy, the results of which may indicate that a certain proportion of the whole population favor or reject the policy.

2.1. Simple Random Sampling

Cochran [8] defines simple random sampling as a method of selecting n units out of a population of size N such that every one of the $\binom{N}{n}$ distinct samples is equally likely to be chosen. He further explains this to mean that at any draw, all the elements in the population that have not already been drawn have an equal chance of being selected.

To draw a simple random sample of size n from a population of size N, the units of the entire population are listed from 1 to N. A unit of the population is selected for inclusion in the sample based on the outcome from a table of random digits or a computer program that produces such a table. That is, a unit is chosen if the selected random number coincides with the list number of that unit. Cochran [8] (pages 18-20), has clearly outlined two different ways of using the table of random digits in selecting a random sample. He also explains sampling with and sampling without replacement.

Let $x_1, x_2, \ldots, x_n$ be a SRS of size n. Then the estimator of the population mean is the sample mean, which is defined as

$$\overline{X}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} x_i \ ,$$

and the variance of $\overline{X}_{SRS}$ for infinite population is defined as

$$Var[\overline{X}_{SRS}] = \frac{\sigma^2}{n},$$

where σ^2 is the population variance. The simple random sample estimator of the population variance is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X}_{SRS})^2$$

2.2. Ranked Set Sampling

The method of ranked set sampling (RSS) as proposed by McIntyre [19] may be summarized as follows. The population under consideration is randomly sampled n times taking n elements at a time, thus yielding n random samples each of size n as shown in Table 2.1. The members of each sample are then ordered with respect to the characteristic under investigation without actual measurement using a cost-free method such as visual inspection (see Table 2.2 on the next page).

TABLE 2.1 n SETS of RANDOM SAMPLES each of SIZE n

X_{11}	X_{12}	...	$X_{1\,n-1}$	X_{1n}
X_{21}	X_{22}	...	$X_{2\,n-1}$	X_{2n}
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
X_{n1}	X_{n2}	...	$X_{n\,n-1}$	X_{nn}

The i^{th} smallest unit from the i^{th} set is then selected for actual measurement. That is the units selected for actual measurement are the smallest unit from the first set, $X_{1(1)}$, the second smallest from the second set, $X_{2(2)}$, up to the largest element, $X_{n(n)}$ from the n^{th} set. Thus, the ranked set sample consists of the diagonal elements of Table 2.2. This is considered as a one cycle ranked set sample. If the n elements of the one cycle does not meet the sample size requirement, the above procedure may be repeated m times to yield a sample of size mn, in which case we have an m-cycle RSS with set size n.

The mean of the selected and measured units from the above procedure was first proposed by McIntyre [20] to be the RSS unbiased estimator of the population mean. That is

$$\overline{X}_{RSS} = \frac{1}{mn} \sum_{j=1}^{m} \sum_{i=1}^{n} X_{i(i)j}$$

where $X_{i(i)j}$ is the i^{th} order statistic from the i^{th} set of the j^{th} cycle. Without any mathematical theory, he argued that this estimator has a smaller variance than the simple random sample estimator.

TABLE 2.2. ORDERED SETS of TABLE 2.1

$X_{1(1)}$	$X_{1(2)}$	...	$X_{1(n-1)}$	$X_{1(n)}$
$X_{2(1)}$	$X_{2(2)}$	...	$X_{2(n-1)}$	$X_{2(n)}$
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
$X_{n(1)}$	$X_{n(2)}$	...	$X_{n(n-1)}$	$X_{n(n)}$

Takahasi and Wakimoto [50] independently proposed the same estimator and mathematically showed it to be unbiased for the population mean μ. Also, they showed that

$$Var(\overline{X}_{RSS}) = \frac{\sigma^2_{(n)}}{nm},$$

where $\sigma^2_{(n)} = \frac{1}{n}\sum_{i=1}^{n}\sigma^2_{n,i}$, the variance of the i^{th} order statistic, and $\sigma^2_{n,i} = E[X_{i(i)j} - E(X_{i(i)j})]^2$. They showed also that $Var(\bar{X}_{RSS})$ is less than $Var(\bar{X}_{SRS})$, the variance of the SRS mean. That is, $\bar{X}_{RSS}$ is more efficient than $\bar{X}_{SRS}$.

Stokes [42] considered the estimation of the population variance, σ^2 using RSS with m cycles. Denoting the i^{th} judgement order statistic in the i^{th} set of the j^{th} cycle by $X_{i[i]j}$, $i=1,2,\ldots,n;\ j=1,2,\ldots,m$ and letting $X_1,\ldots,X_{mn}$ denote a SRS of mn elements chosen from the same population with variance σ^2, the following estimators of σ^2 were defined;

$$s^2 = \sum_{j=1}^{mn}\left(X_j - \bar{X}\right)^2 \Big/ (mn-1) \qquad \text{and} \qquad \hat{\sigma}^2 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(X_{i[i]j} - \bar{X}_{RSS}\right)^2 \Big/ (mn-1), \qquad \text{where}$$

$$\bar{X}_{RSS} = \sum_{j=1}^{m}\sum_{i=1}^{n} X_{i[i]j} / mn.$$

s^2 and $\hat{\sigma}^2$ are respectively the SRS and RSS estimates of σ^2. It was noted that $\hat{\sigma}^2$ is a biased estimator of σ^2. It was however shown to be approximately unbiased for large values of either m or n or both. However, it is more practical to limit the largeness to m as in the case of large n, ranking becomes difficult thereby sacrificing the fundamental requirement of the MTW procedure. She proved that the relative precision

$$RP = \frac{Var(s^2)}{MSE(\hat{\sigma}^2)} \geq 1 \quad as \quad mn \to \infty.$$

In her 1995 paper, Stokes [45] gives the following results, which we shall be using throughout this thesis. Let $\{X_{i(i)j}, i = 1,2,\ldots,n; j = 1,2,\ldots,m\}$ be a RSS of set size n and cycle size m from a distribution with cumulative distribution function $F\left(\frac{x-\mu}{\sigma}\right)$ and probability density function $\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$. Using this pdf and assuming the validity of the usual regularity conditions, the loglikelihood function of the RSS is

$$L = K - mn\ln\sigma + \sum_{j=1}^{m}\sum_{i=1}^{n} \ln f(Z_{i(i)j}) + \sum_{j=1}^{m}\sum_{i=1}^{n} (i-1)\ln F(Z_{i(i)j}) + \sum_{j=1}^{m}\sum_{i=1}^{n} (n-i)\ln[1 - F(Z_{i(i)j})]$$

where $Z_{i(i)j} = \frac{X_{i(i)j} - \mu}{\sigma}$.

The maximum likelihood estimate of μ, $\hat{\mu}^{*}_{ML}$ is the solution of the equation

$$-\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{f'(Z_{i(i)j})}{f(Z_{i(i)j})} + \sum_{j=1}^{m}\sum_{i=1}^{n} (i-1)\frac{f(Z_{i(i)j})}{F(Z_{i(i)j})} + \sum_{j=1}^{m}\sum_{i=1}^{n} (n-i)\frac{f(Z_{i(i)j})}{[1 - F(Z_{i(i)j})]} = 0$$

and the Fisher information for μ is

$$I^{*}_{mn}(\mu) = \frac{mn}{\sigma^2} E\left\{\frac{f'(Z_r)}{f(Z_r)}\right\}^2 + \frac{mn(n-1)}{\sigma^2} E\left\{\frac{f^2(Z_r)}{F(Z_r)[1 - F(Z_r)]}\right\}.$$

For a simple random sample of size mn from the same population, the maximum likelihood estimator of μ, $\hat{\mu}_{ML}$ is the solution of the equation

$$\sum_{r=1}^{mn} \frac{f'(Z_r)}{f(Z_r)} = 0,$$

and the Fisher information for μ from the SRS is

$$I_{mn}(\mu) = \frac{mn}{\sigma^2} E\left\{\frac{f'(Z_r)}{f(Z_r)}\right\}.$$

Thus the asymptotic relative precision of RSS relative to SRS is

$$\begin{aligned}\lim_{m\to\infty} RP(\hat{\mu}^{*}_{ML}, \hat{\mu}_{ML}) &= \frac{I^{*}_{mn}(\mu)}{I_{mn}(\mu)} \\ &= 1 + (n-1)E\left\{\frac{f^2(Z_r)}{F(Z_r)[1 - F(Z_r)]}\right\} \Big/ E\left\{\frac{f'(Z_r)}{f(Z_r)}\right\}^2\end{aligned}$$

Similar results for σ are shown in Stokes [45], who clearly demonstrates the dominance of the RSS estimators over the SRS estimators.

2.3. Median Ranked Set Sampling

As proposed by Muttlak [25], this method depends on whether the set size is even or odd. For odd set sizes, the median value is selected from each of the n ordered sets (See Table 2.2). However, for even set sizes, the $(n / 2)^{th}$ smallest element is chosen from the first $n / 2$ ordered sets, while the $((n + 2) / 2))^{th}$ smallest unit is chosen from each of the remaining $n / 2$ sets.

In set notation, the MRSS is the set $\left\{X_{i(\frac{n+1}{2})j},\quad i=1,2,\dots,n; j=1,2,\dots,m\right\}$ for odd set, where $X_{i(\frac{n+1}{2})j}$ is the i^{th} median from the i^{th} set of the j^{th} cycle. For even set sizes, the MRSS in set notation is

$$\left\{X_{i(\frac{n}{2})j},\quad i=1,2,\dots,n/2; j=1,2,\dots,m\right\}\bigcup\left\{X_{i(\frac{n+2}{2})j},\quad i=\frac{n}{2}+1,\dots,n; j=1,2,\dots,m\right\}$$

where $X_{i(\frac{n}{2})j}$ and $X_{i(\frac{n+2}{2})j}$ are respectively, the $(n / 2)^{th}$ and the $((n + 2) / 2))^{th}$ smallest units from the i^{th} set of the j^{th} cycle.

Assuming one cycle (i.e. $m = 1$), the suggested estimators (Muttlak [25]) of the population mean μ for odd and even sample sizes respectively are

$$\overline{X}_{MRSS1} = \frac{1}{n}\sum_{i=1}^{n} X_{i((n+1)/2)}$$

where $X_{i((n+1)/2}$ is the $((n+1)/2)^{th}$ smallest rank in the i^{th} set and

$$\overline{X}_{MRSS2} = \frac{1}{n}\left(\sum_{i=1}^{L} X_{i(n/2)} + \sum_{i=L+1}^{n} X_{i((n+2)/2)}\right)$$

where $X_{i((n/2))}$ and $X_{i((n+2)/2)}$ are respectively, the $(n/2)^{th}$ and $((n+2)/2)^{th}$ smallest ranked unit in the i^{th} set and $L=n/2$. It was shown that each of the variances of these estimators, $Var(\overline{X}_{MRSS1})$ and $Var(\overline{X}_{MRSS2})$ is less than the variance for the simple random sample mean, $Var(\overline{X}_{SRS})$, if the underlying distribution is symmetric about the population mean μ. In cases of asymmetric distributions, the mean square error of the first estimator is defined as

$$MSE(\overline{X}_{MRSS1}) = Var(\overline{X}_{MRSS1}) + (bias)^2,$$

where $bias = \mu - E(\overline{X}_{MRSS1})$. This is shown to be less than $Var(\overline{X}_{SRS})$. A similar result holds for the second estimator.

2.4. Extreme Ranked Set Sampling

The method of ERSS as studied by Samawi et al. [36] draws n times, a random sample of size n from a population under consideration. For even set size n, the largest and smallest units are alternately taken from the first to the n^{th} random sample. The resulting sample of $n/2$ each of smallest and the largest observations forms the extreme ranked set sample (ERSS). On the other hand, if n is odd, the largest and smallest units are selected from the first random sample to the $(n-1)^{st}$ random sample. From the n^{th} random sample, either the mean of the largest and smallest unit is chosen or the median of the whole set. In their paper, Samawi et al. [36] highlight the practicability of the first alternative as opposed to the second, arguing that taking the median of the set entails complete ranking of the set whereas taking the average of the smallest and the largest unit does not. We would however be using the median since we are interested in a set of independent observations.

Under each of ERSS, the estimators of the population mean are unbiased for symmetric distributions and biased otherwise. Some dominance of the ERSS over SRS has been demonstrated.

CHAPTER 3

MAXIMUM LIKELIHOOD ESTIMATION USING MEDIAN RANKED SET SAMPLING

In this chapter, we consider the family of distributions with cumulative distribution function of the form $F\left(\frac{x-\mu}{\sigma}\right)$ and probability density function $\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$. That is a family of distributions with the specified functions with one or none of μ and σ known. In each case, we will consider the asymptotic relative precision of the maximum likelihood estimator (MLE) of the unknown parameter under median ranked set sampling (MRSS) and compare our results to those obtained under ranked set sampling (Stokes [45]).

Suppose $X_1, X_2, \ldots, X_n$ are n independent variates from a population with cumulative distribution function (cdf) $F(x)$ and probability distribution function (pdf) $f(x)$. Let $F_r(x)$

and $f_r(x)$ denote the cdf and pdf respectively of the r^{th} order statistic. It can be shown (David [9]) that

$$F_r(x)=\sum_{k=r}^{n}\binom{n}{k}F^k(x)[1-F(x)]^{n-k} \tag{3.1}$$

and

$$f_r(x)=\frac{1}{B(r,n-r+1)}F^{r-1}(x)[1-F(x)]^{n-r}f(x), \tag{3.2}$$

where B(.,.) denotes the beta function. In using the MRSS, we would be interested in the pdfs of the $(n/2)^{th}$ and $((n+2)/2)^{th}$ order statistics when n is even, and in that of the $((n+1)/2)^{th}$ order statistic when n is odd. Direct substitution into equation (3.2) gives the following pdfs

$$f_{\frac{n}{2}}(x)=\frac{1}{B\left(\frac{n}{2},\frac{n}{2}+1\right)}F^{\frac{n-2}{2}}(x)[1-F(x)]^{\frac{n}{2}}f(x), \tag{3.3}$$

$$f_{\frac{n+2}{2}}(x)=\frac{1}{B\left(\frac{n+2}{2},\frac{n}{2}\right)}F^{\frac{n}{2}}(x)[1-F(x)]^{\frac{n-2}{2}}f(x), \tag{3.4}$$

$$f_{\frac{n+1}{2}}(x)=\frac{1}{B\left(\frac{n+1}{2},\frac{n+1}{2}\right)}F^{\frac{n-1}{2}}(x)[1-F(x)]^{\frac{n-1}{2}}f(x). \tag{3.5}$$

Now suppose we have a MRSS with set size n and cycle size m from a population with cdf $F\left(\frac{x-\mu}{\sigma}\right)$ and pdf $\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$. We wish to estimate the parameters μ and σ using the method of maximum likelihood.

Let $Z=\frac{X-\mu}{\sigma}$, then $F\left(\frac{X-\mu}{\sigma}\right)=F(Z)$ and $f\left(\frac{X-\mu}{\sigma}\right)=f(Z)$.

Before we proceed, we state the following derivatives, which we shall be using in the process.

(a) Derivatives with respect to μ

(i) $\dfrac{\partial f(Z)}{\partial \mu} = -\dfrac{1}{\sigma} f'(Z)$,

(ii) $\dfrac{\partial F(Z)}{\partial \mu} = -\dfrac{1}{\sigma} f(Z)$,

(iii) $\dfrac{\partial}{\partial \mu} \ln f(Z) = -\dfrac{1}{\sigma}\dfrac{f'(Z)}{f(Z)}$,

(iv) $\dfrac{\partial}{\partial \mu} \ln F(Z) = -\dfrac{1}{\sigma}\dfrac{f(Z)}{F(Z)}$,

(v) $\dfrac{\partial}{\partial \mu} \ln[1-F(Z)] = \dfrac{1}{\sigma}\dfrac{f(Z)}{[1-F(Z)]}$,

(vi) $\dfrac{\partial}{\partial \mu}\dfrac{f(Z)}{F(Z)} = -\dfrac{1}{\sigma}\left[\dfrac{f'(Z)}{F(Z)} - \left(\dfrac{f(Z)}{F(Z)}\right)^2\right]$,

(vii) $\dfrac{\partial}{\partial \mu}\dfrac{f(Z)}{1-F(Z)} = -\dfrac{1}{\sigma}\left[\dfrac{f'(Z)}{1-F(Z)} + \left(\dfrac{f(Z)}{1-F(Z)}\right)^2\right]$,

(viii) $\dfrac{\partial}{\partial \mu}\dfrac{f'(Z)}{f(Z)} = -\dfrac{1}{\sigma}\left[\dfrac{f''(Z)}{f(Z)} - \left(\dfrac{f'(Z)}{f(Z)}\right)^2\right]$.

(b) Derivatives with respect to σ

(i) $\dfrac{\partial f(Z)}{\partial \sigma} = -\dfrac{Z}{\sigma} f'(Z)$,

(ii) $\dfrac{\partial F(Z)}{\partial \sigma} = -\dfrac{Z}{\sigma} f(Z)$,

(iii) $\dfrac{\partial}{\partial \sigma} \ln f(Z) = -\dfrac{Z}{\sigma}\dfrac{f'(Z)}{f(Z)}$,

(iv) $\dfrac{\partial}{\partial \sigma} \ln F(Z) = -\dfrac{Z}{\sigma}\dfrac{f(Z)}{F(Z)}$,

(v) $\frac{\partial}{\partial\sigma}\ln(1-F(Z))=\frac{Z}{\sigma}\frac{f(Z)}{(1-F(Z))}$,

(vi) $\frac{\partial}{\partial\sigma}\left(\frac{Zf(Z)}{1-F(Z)}\right)=-\frac{1}{\sigma}\left[\frac{Z^2f'(Z)+Zf(Z)}{1-F(Z)}+\left(\frac{Zf(Z)}{1-F(Z)}\right)^2\right]$,

(vii) $\frac{\partial}{\partial\sigma}\left(\frac{Zf(Z)}{F(Z)}\right)=-\frac{1}{\sigma}\left[\frac{Z^2f'(Z)+Zf(Z)}{F(Z)}-\left(\frac{Zf(Z)}{F(Z)}\right)^2\right]$,

(viii) $\frac{\partial}{\partial\sigma}\left(\frac{Zf'(Z)}{f(Z)}\right)=-\frac{1}{\sigma}\left[\frac{Z^2f''(Z)+Zf'(Z)}{f(Z)}-\left(\frac{Zf'(Z)}{f(Z)}\right)^2\right]$.

We will now proceed with the maximum likelihood estimation under the two cases of the MRSS, i.e. with odd and even set sizes.

3.1. The MLE with Odd Set Size

Suppose $\{X_{i(p)j}, i=1,2,...,n;\ j=1,2,...,m\}$ is a MRSS from a population with distribution function F and probability density function f, where $p = (n + 1)/2$. Suppose further that the distribution is of the location-scale type. Further, let $Z_{i(p)j}=\frac{X_{i(p)j}-\mu}{\sigma}$.

Then the $Z_{i(p)j}$'s are independent and identically distributed with pdf

$$f_p(z)=\frac{1}{B(p,p)}F^{(p-1)}(z)[1-F(z)]^{(p-1)}\frac{1}{\sigma}f(z), \tag{3.6}$$

where $p=\frac{n+1}{2}$.

Thus, the likelihood function of the MRSS is

$$l_{MRSS1}=\prod_{j=1}^{m}\prod_{i=1}^{j}f_p(Z_{i(p)j}),$$

and the loglikelihood function is

$$L_{MRSS1} = \ln \prod_{j=1}^{m}\prod_{i=1}^{n} f_p(Z_{i(p)j}) = \sum_{j=1}^{m} \ln \prod_{i=1}^{n} f_p(Z_{i(p)j}) = \sum_{j=1}^{m}\sum_{i=1}^{n} \ln f_p(Z_{i(p)j})$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{n}\left\{K - \ln\sigma + \left(\frac{n-1}{2}\right)\left[\ln F(Z_{i(p)j}) + \ln\left(1 - F(Z_{i(p)j})\right)\right]\right\} + \sum_{j=1}^{m}\sum_{i=1}^{n} \ln f(Z_{i(p)j})$$

$$= K_1 - mn\ln\sigma + \left(\frac{n-1}{2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\ln F(Z_{i(p)j}) + [1 - F(Z_{i(p)j})]\right\} + \sum_{j=1}^{m}\sum_{i=1}^{n} \ln f(Z_{i(p)j}), \tag{3.7}$$

where $\frac{n-1}{2} = p-1$.

Now suppose that σ is known. Then we estimate μ by differentiating equation (3.7) with respect to μ, setting the result equal to zero and solving for μ. Differentiating with respect to μ, we get

$$\frac{\partial L_{MRSS1}}{\partial \mu} = \left(\frac{n-1}{2\sigma}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{f(Z_{i(p)j})}{1-F(Z_{i(p)j})} - \frac{f(Z_{i(p)j})}{F(Z_{i(p)j})}\right\} - \frac{1}{\sigma}\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{f'(Z_{i(p)j})}{f(Z_{i(p)j})}. \tag{3.8}$$

Thus, the maximum likelihood estimator $\hat{\mu}_{mle1}$ of μ is the solution of the equation

$$\frac{n-1}{2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{f(Z_{i(p)j})}{1-F(Z_{i(p)j})} - \frac{f(Z_{i(p)j})}{F(Z_{i(p)j})}\right\} - \sum_{j=1}^{m}\sum_{i=1}^{n}\frac{f'(Z_{i(p)j})}{f(Z_{i(p)j})} = 0. \tag{3.9}$$

To find the Fisher information, we assume the validity of the regularity conditions and then differentiate equation (3.8) with respect to μ to obtain

$$\frac{\partial^2 L_{MRSS1}}{\partial \mu^2} = \left(\frac{n-1}{2\sigma^2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{f'(Z_{i(p)j})}{F(Z_{i(p)j})} - \left(\frac{f(Z_{i(p)j})}{F(Z_{i(p)j})}\right)^2 - \left[\frac{f'(Z_{i(p)j})}{1-F(Z_{i(p)j})} + \left(\frac{f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right)^2\right]\right\}$$

$$+\frac{1}{\sigma^2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left[\frac{f''(Z_{i(p)j})}{f(Z_{i(p)j})} - \left(\frac{f'(Z_{i(p)j})}{f(Z_{i(p)j})}\right)^2\right].$$

The Fisher information about μ from the MRSS is

$$I_{mn1}(\mu) = -E\left\{\frac{\partial^2 L_{MRSS1}}{\partial \mu^2}\right\}$$

$$= -\frac{mn(n-1)}{2\sigma^2} E \sum_{j=1}^{m} \sum_{i=1}^{n} \left\{ \frac{f'(Z_{i(p)j})}{F(Z_{i(p)j})} - \left(\frac{f(Z_{i(p)j})}{F(Z_{i(p)j})} \right)^2 - \frac{f'(Z_{i(p)j})}{1-F(Z_{i(p)j})} - \left(\frac{f(Z_{i(p)j})}{1-F(Z_{i(p)j})} \right)^2 \right\}$$

$$+ \frac{mn}{\sigma^2} E \sum_{j=1}^{m} \sum_{i=1}^{n} \left\{ \left(\frac{f'(Z_{i(p)j})}{f(Z_{i(p)j})} \right)^2 - \frac{f''(Z_{i(p)j})}{f(Z_{i(p)j})} \right\},$$

which may be simplified to give

$$I_{mn1}(\mu) = -\frac{mn(n-1)}{2\sigma^2} E\left[\left\{ \frac{f'(z)}{F(z)} - \left[\frac{f(z)}{F(z)} \right]^2 - \frac{f'(z)}{1-F(z)} - \left[\frac{f(z)}{1-F(z)} \right]^2 \right\} g(z) \right]$$

$$\frac{mn}{\sigma^2} E\left\{ \left[\left(\frac{f'(z)}{f(z)} \right)^2 - \frac{f''(z)}{f(z)} \right] g(z) \right\}$$

$$= \frac{mn}{\sigma^2} E\left\{ \left[\left(\frac{f'(z)}{f(z)} \right)^2 - \frac{f''(z)}{f(Z_r)} \right] g(z) \right\}$$

$$+ \frac{mn(n-1)}{2\sigma^2} E\left[\left\{ \frac{[2F(z)-1]f'(z)}{F(z)[1-F(z)]} + \frac{[F^2(z)+(1-F(z))^2]f^2(z)}{[1-F(z)]^2 F^2(z)} \right\} g(z) \right], \tag{3.10}$$

where

$$g(z) = \frac{1}{B(p,p)} F^{p-1}(z)[1-F(z)]^{p-1} \tag{3.11}$$

To prove the result in equation (3.10), we will only show that

$$E\left\{ \sum_{j=1}^{n} \sum_{i=1}^{n} \left\{ \frac{f'(Z_{i(p)j})}{F(Z_{i(p)j})} - \left[\frac{f(Z_{i(p)j})}{F(Z_{i(p)j})} \right]^2 \right\} \right\} = mnE\left[\left\{ \frac{f'(z)}{F(z)} - \left[\frac{f(z)}{F(z)} \right]^2 \right\} g(z) \right],$$

since the other parts can similarly be shown.

Proof:

Assuming that the regularity conditions are satisfied, let

$h(z) = \frac{f'(z)}{F(z)} - \left[\frac{f(z)}{F(z)}\right]^2$. Then the required expectation is

$$\int_{-\infty}^{\infty}\sum_{j=1}^{m}\sum_{i=1}^{n} h(z)f_p(z)dz=\sum_{j=1}^{m}\sum_{i=1}^{n}\int_{-\infty}^{\infty} h(z)\frac{1}{B(p,p)}F^{p-1}(z)[1-F(z)]^{p-1}f(z)dz$$
$$=mnE\left[h(z)\frac{1}{B(p,p)}F^{p-1}(z)[1-F(z)]^{p-1}\right]$$
$$=mnE[h(z)g(z)].$$

As in Stokes [45] the loglikelihood function of a simple random sample of size nm is

$$L=-mn\ln\sigma+\sum_{i=1}^{mn}\ln f(Z_i), \tag{3.12}$$

where $Z_i=\dfrac{X_i-\mu}{\sigma}$. The maximum likelihood estimator $\hat{\mu}_{ML}$, from a random sample of size nm, is the solution of the equation

$$\sum_{i=1}^{mn}\frac{f'(Z_r)}{f(Z_r)}=0, \tag{3.13}$$

and the Fisher Information for μ is

$$I_{mn}(\mu)=\frac{mn}{\sigma^2}E\left\{\frac{f'(z)}{f(z)}\right\}^2, \tag{3.14}$$

provided $E\left\{\dfrac{\partial\ln f}{\partial\mu}\right\}=0$.

Following Stokes [45], we define the asymptotic relative precision of MRSS with respect to SRS for estimating μ as

$$\lim_{m\to\infty} RP(\hat{\mu}_{mle1},\hat{\mu}_{ML})=\frac{I_{mn1}(\mu)}{I_{mn}(\mu)}. \tag{3.15}$$

Now, suppose that σ is to be estimated and μ is known. Differentiating equation (3.7) with respect to σ, we get

$$\frac{\partial L_{MRSS1}}{\partial\sigma}=-\frac{mn}{\sigma}+\left(\frac{n-1}{2\sigma}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{-\frac{Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}+\frac{Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right\}-\frac{1}{\sigma}\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}.$$

Therefore, the maximum likelihood estimator, $\hat{\sigma}_{mle1}$ of σ is the solution of

$$-mn+\frac{n-1}{2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{-\frac{Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}+\frac{Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right\}-\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}=0. \tag{3.16}$$

Further, to find the Fisher information for σ from the MRSS, we need to take the second derivative with respect to σ, again assuming that the regularity conditions are satisfied, to get

$$\begin{aligned}
\frac{\partial^2 L_{MRSS1}}{\partial\sigma^2}=&\frac{mn}{\sigma^2}-\left(\frac{n-1}{2\sigma^2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{-\frac{Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}+\frac{Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right\}\\
&-\left(\frac{n-1}{2\sigma^2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{-\frac{Z^2_{i(p)j}f'(Z_{i(p)j})+Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}+\left(\frac{Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}\right)^2\right.\\
&\left.+\left[\frac{Z^2_{i(p)j}f'(Z_{i(p)j})+Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}+\left(\frac{Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right)^2\right]\right\}\\
&+\frac{1}{\sigma^2}\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}\\
&+\frac{1}{\sigma^2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{Z^2_{i(p)j}f''(Z_{i(p)j})+Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}-\left(\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}\right)^2\right\},
\end{aligned}$$

$$\frac{\partial^2 L_{MRSSI}}{\partial\sigma^2}=\frac{mn}{\sigma^2}-\left(\frac{n-1}{2\sigma^2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{-\frac{Z^2_{i(p)j}f'(Z_{i(p)j})+2Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}+\left(\frac{Z_{i(p)j}f(Z_{i(p)j})}{F(Z_{i(p)j})}\right)^2\right.$$
$$\left.+\left[\frac{Z^2_{i(p)j}f'(Z_{i(p)j})+2Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}+\left(\frac{Z_{i(p)j}f(Z_{i(p)j})}{1-F(Z_{i(p)j})}\right)^2\right]\right\}$$
$$+\frac{1}{\sigma^2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{Z^2_{i(p)j}f''(Z_{i(p)j})+2Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}-\left(\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}\right)^2\right\}$$
$$=\frac{mn}{\sigma^2}-\left(\frac{n-1}{2\sigma^2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{[2F(Z_{i(p)j})-1][Z^2_{i(p)j}f'(Z_{i(p)j})+2Z_{i(p)j}f(Z_{i(p)j})]}{[1-F(Z_{i(p)j})]F(Z_{i(p)j})}\right.$$
$$\left.+\frac{[F^2(Z_{i(p)j})+(1-F(Z_{i(p)j}))^2]Z^2_{i(p)j}f(Z_{i(p)j})}{F^2(Z_{i(p)j})[1-F(Z_{i(p)j})]^2}\right\}$$
$$+\frac{1}{\sigma^2}\sum_{j=1}^{m}\sum_{i=1}^{n}\left\{\frac{Z^2_{i(p)j}f''(Z_{i(p)j})+2Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}-\left(\frac{Z_{i(p)j}f'(Z_{i(p)j})}{f(Z_{i(p)j})}\right)^2\right\}.$$

Taking the negative expectation of this gives the Fisher information

$$I_{mn1}(\sigma)=-\frac{mn}{\sigma^2}+\frac{mn(n-1)}{2\sigma^2}E\left[\left\{-\frac{z^2f'(z)+2zf(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right.\right.$$
$$\left.\left.+\left(\frac{z^2f'(z)+2zf(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)\right\}g(z)\right]$$
$$-\frac{mn}{\sigma^2}E\left\{\left[\frac{z^2f''(z)+2zf'(z)}{f(z)}-\left(\frac{zf'(z)}{f(z)}\right)^2\right]g(z)\right\}$$

$$=\frac{mn}{\sigma^2}E\left\{\left[\left(\frac{zf'(z)}{f(z)}\right)^2-\frac{z^2f''(z)+2zf'(z)}{f(z)}\right]g(z)-1\right\}$$
$$+\frac{mn(n-1)}{2\sigma^2}E\left[\left\{\frac{[2F(z)-1][z^2f'(z)+2zf(z)]}{[1-F(z)]F(z)}\right.\right. \tag{3.17}$$
$$\left.\left.+\frac{[F^2(z)+(1-F(z))^2]z^2f^2(z)}{F^2(z)[1-F(z)]^2}\right\}g(z)\right]$$

where $g(z)$ is as defined in equation (3.11). It can easily be shown (Stokes [45]) that from a random sample of size mn, the likelihood estimator of σ is the solution of the equation

$$mn+\sum_{r=1}^{mn}\frac{Z_r f'(Z_r)}{f(Z_r)}=0,$$

and the Fisher information is

$$I_{mn}(\sigma)=\frac{mn}{\sigma^2}E\left\{\frac{zf'(z)}{f(z)}\right\}.$$

Thus, we define the asymptotic relative precision of MRSS with respect to SRS for the estimation of σ as

$$\lim_{m\to\infty} RP\left(\hat{\sigma}_{ML_2},\hat{\sigma}_{ML}\right)=\frac{I_{mn2}(\sigma)}{I_{mn}(\sigma)}. \tag{3.18}$$

3.2. MLE with Even Set Size

Let $q = n / 2$ and $\left\{\{X_{i(q)j}\}_{i=1}^{q}\bigcup\{X_{i(q+1)j}\}_{i=q+1}^{n};\quad j=1,2,\dots,m\right\}$ be a MRSS, where n and m are the set size and the cycle size respectively. Let $Z_{i(q)j}$ and $Z_{i(q+1)j}$ be the corresponding standardized order statistics analogous to $Z_{i(p)j}$ in the case of odd set size. Then their respective pdf's are

$$f_q(z)=\frac{1}{\sigma\, B(q,q+1)}F^{(q-1)}(z)[1-F(z)]^q f(z), \tag{3.19}$$

and

$$f_{q+1}(z)=\frac{1}{\sigma\, B(q+1,q)}F^q(z)[1-F(z)]^{(q-1)} f(z). \tag{3.20}$$

Thus, the likelihood function is

$$l_{MRSS2} = \prod_{j=1}^{m}\left[\prod_{i=1}^{q} f_q(Z_{i(q)j})\prod_{i=q+1}^{n} f_{q+1}(Z_{i(q+1)j})\right].$$

The loglikelihood function is

$$\begin{aligned}
L_{MRSS2} &= \ln\prod_{j=1}^{m}\left[\prod_{i=1}^{q} f_q(Z_{i(q)j})\prod_{i=q+1}^{n} f_{q+1}(Z_{i(q+1)j})\right] \\
&= \sum_{j=1}^{m}\ln\left[\prod_{i=1}^{q} f_q(Z_{i(q)j})\prod_{i=q+1}^{n} f_{q+1}(Z_{i(q+1)j})\right] \\
&= \sum_{j=1}^{m}\left[\ln\prod_{i=1}^{q} f_q(Z_{i(q)j})+\ln\prod_{i=q+1}^{n} f_{q+1}(Z_{i(q+1)j})\right] \\
&= \sum_{j=1}^{m}\left[\sum_{i=1}^{q}\ln f_q(Z_{i(q)j})+\sum_{i=q+1}^{n}\ln f_{q+1}(Z_{i(q+1)j})\right].
\end{aligned}$$

Now, substituting equation (3.19) and equation (3.20), we obtain

$$\begin{aligned}
L_{MRSS2} = K_2 - mn\ln\sigma + \sum_{j=1}^{m}\Bigg\{&\sum_{i=1}^{q}\left[\frac{n-2}{2}\ln F(Z_{i(q)j})+\frac{n}{2}\ln[1-F(Z_{i(q)j})]+\ln f(Z_{i(q)j})\right] \\
&+\sum_{i=q+1}^{n}\left[\frac{n}{2}\ln F(Z_{i(q+1)j})+\frac{n-2}{2}\ln[1-F(Z_{i(q+1)j})]+\ln f(Z_{i(q+1)j})\right]\Bigg\}.
\end{aligned} \tag{3.21}$$

Again assuming that the regularity conditions hold and differentiating equation (3.21) with respect to μ, we get

$$\begin{aligned}
\frac{\partial L_{MRSS2}}{\partial\mu} = \frac{1}{\sigma}\sum_{j=1}^{m}\Bigg\{&\sum_{i=1}^{q}\left[-\frac{n-2}{2}\frac{f(Z_{i(q)j})}{F(Z_{i(q)j})}+\frac{n}{2}\frac{f(Z_{i(q)j})}{[1-F(Z_{i(q)j})]}-\frac{f'(Z_{i(q)j})}{f(Z_{i(q)j})}\right] \\
&+\sum_{i=q+1}^{n}\left[-\frac{n}{2}\frac{f(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}+\frac{n-2}{2}\frac{f(Z_{i(q+1)j})}{[1-F(Z_{i(q+1)j})]}-\frac{f'(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}\right]\Bigg\}.
\end{aligned} \tag{3.22}$$

Setting this equal to zero and solving for μ yields the maximum likelihood estimator, $\hat{\mu}_{mle2}$. To find the Fisher information, we differentiate equation (3.22) with respect to μ to get

$$\frac{\partial^2 L_{MRSS2}}{\partial\mu^2}=-\frac{1}{\sigma^2}\sum_{j=1}^{m}\left\{\sum_{i=1}^{q}\left[-\left(\frac{n-2}{2}\right)\left(\frac{f'(Z_{i(q)j})}{F(Z_{i(q)j})}-\left(\frac{f(Z_{i(q)j})}{F(Z_{i(q)j})}\right)^2\right)\right.\right.$$
$$+\left(\frac{n}{2}\right)\left(\frac{f'(Z_{i(q)j})}{1-F(Z_{i(q)j})}+\left(\frac{f(Z_{i(q)j})}{1-F(Z_{i(q)j})}\right)^2\right)$$
$$\left.-\left(\frac{f''(Z_{i(q)j})}{f(Z_{i(q)j})}-\left(\frac{f'(Z_{i(q)j})}{f(Z_{i(q)j})}\right)^2\right)\right]$$
$$+\sum_{i=q+1}^{n}\left[-\left(\frac{n}{2}\right)\left(\frac{f'(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}-\left(\frac{f(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}\right)^2\right)\right.$$
$$+\left(\frac{n-2}{2}\right)\left(\frac{f'(Z_{i(q+1)j})}{1-F(Z_{i(q+1)j})}+\left(\frac{f(Z_{i(q+1)j})}{1-F(Z_{i(q+1)j})}\right)^2\right)$$
$$\left.\left.-\left(\frac{f''(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}-\left(\frac{f'(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}\right)^2\right)\right]\right\}.$$

Finally, we find the negative expectation and obtain the Fisher information as

$$I_{mn2}(\mu)=\frac{1}{\sigma^2}E\left[\sum_{j=1}^{m}\left\{\sum_{i=1}^{q}\left[-\left(\frac{n-2}{2}\right)\left(\frac{f'(Z_{i(q)j})}{F(Z_{i(q)j})}-\left(\frac{f(Z_{i(q)j})}{F(Z_{i(q)j})}\right)^2\right)\right.\right.\right.$$
$$+\left(\frac{n}{2}\right)\left(\frac{f'(Z_{i(q)j})}{1-F(Z_{i(q)j})}+\left(\frac{f(Z_{i(q)j})}{1-F(Z_{i(q)j})}\right)^2\right)$$
$$\left.-\left(\frac{f''(Z_{i(q)j})}{f(Z_{i(q)j})}-\left(\frac{f'(Z_{i(q)j})}{f(Z_{i(q)j})}\right)^2\right)\right]$$
$$+\sum_{i=q+1}^{n}\left[-\left(\frac{n}{2}\right)\left(\frac{f'(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}-\left(\frac{f(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}\right)^2\right)\right.$$
$$+\left(\frac{n-2}{2}\right)\left(\frac{f'(Z_{i(q+1)j})}{1-F(Z_{i(q+1)j})}+\left(\frac{f(Z_{i(q+1)j})}{1-F(Z_{i(q+1)j})}\right)^2\right)$$
$$\left.\left.\left.-\left(\frac{f''(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}-\left(\frac{f'(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}\right)^2\right)\right]\right\}\right].$$

Let

$$g_1(z)=\frac{1}{B(q,q+1)}F^{q-1}(z)[1-F(z)]^{q}, \tag{3.23}$$

and

$$g_2(z)=\frac{1}{B(q,q+1)}F^q(z)[1-F(z)]^{q-1} \ , \tag{3.24}$$

where $q=n/2$. Then the Fisher information can easily be shown (as in section 3.1) to be given by

$$\begin{aligned} I_{mn2}(\mu) &= \frac{mn}{2\sigma^2}E\left\{(g_1(z)+g_2(z))\left[\left(\frac{f'(z)}{f(z)}\right)^2-\frac{f''(z)}{f(z)}\right]\right\} \\ &+\frac{mn}{4\sigma^2}E\left\{n\left[\left(\frac{f'(z)}{1-F(z)}+\left(\frac{f(z)}{1-F(z)}\right)^2\right)g_1(z)-\left(\frac{f'(z)}{F(z)}-\left(\frac{f(z)}{F(Z_r)}\right)^2\right)g_2(z)\right]\right. \\ &\left.+(n-2)\left[\left(\frac{f'(z)}{1-F(z)}+\left(\frac{f(z)}{1-F(z)}\right)^2\right)g_2(z)-\left(\frac{f'(z)}{F(z)}-\left(\frac{f(z)}{F(z)}\right)^2\right)g_1(z)\right]\right\}. \end{aligned} \tag{3.25}$$

Thus, the corresponding asymptotic relative precision is

$$\lim_{m\to\infty} RP(\hat{\mu}_{mle2},\hat{\mu}_{ML})=\frac{I_{mn2}(\mu)}{I_{mn}(\mu)}.$$

To estimate σ when μ is known, we differentiate equation (3.21) with respect to σ to get

$$\begin{aligned} \frac{\partial L_{MRSS2}}{\partial\sigma} = -\frac{mn}{\sigma}-\frac{1}{\sigma}\sum_{j=1}^{m}&\left\{\sum_{i=1}^{q}\left[\left(\frac{n-2}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}-\left(\frac{n}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}+\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right]\right. \\ &\left.+\sum_{i=q+1}^{n}\left[\left(\frac{n}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}-\left(\frac{n-2}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}+\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right]\right\}. \end{aligned}$$

Thus, the MLE for σ, $\hat{\sigma}_{mle2}$ is the solution of the equation

$$\begin{aligned} mn+\sum_{j=1}^{m}&\left\{\sum_{i=1}^{q}\left[(n-2)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}-(n)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}+\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right]\right. \\ &\left.+\sum_{i=q+1}^{n}\left[(n)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}-(n-2)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}+\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right]\right\}=0. \end{aligned} \tag{3.26}$$

To find the Fisher information for σ when μ is known, we take the second derivative with respect to σ to get

$$
\begin{aligned}
\frac{\partial^2 L_{MRSS2}}{\partial\sigma^2} = \frac{mn}{\sigma^2} + \frac{1}{\sigma^2}\sum_{j=1}^{m}\Bigg\{ & \sum_{i=1}^{q}\Bigg[\left(\frac{n-2}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})} - \left(\frac{n}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})} + \frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\Bigg] \\
& + \sum_{i=q+1}^{n}\Bigg[\left(\frac{n}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q+1)j})} - \left(\frac{n-2}{2}\right)\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q+1)j})} + \frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q+1)j})}\Bigg]\Bigg\} \\
+ \frac{1}{\sigma^2}\sum_{j=1}^{m}\Bigg\{ & \sum_{i=1}^{q}\Bigg[\left(\frac{n-2}{2}\right)\left(\frac{Z_{i(q)j}^2 f'(Z_{i(q)j}) + Z_{i(q)j}f(Z_{i(q)j})}{F(Z_{i(q)j})} - \left(\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}\right)^2\right) \\
& - \frac{n}{2}\left(\frac{Z_{i(q)j}^2 f'(Z_{i(q)j}) + Z_{i(q)j}f(Z_{i(q)j})}{1-F(Z_{i(q)j})} + \left(\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}\right)^2\right) \\
& + \frac{Z_{i(q)j}^2 f''(Z_{i(q)j}) + Z_{i(q)j}f'(Z_{i(q)j})}{f(Z_{i(q)j})} - \left(\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right)^2\Bigg] \\
& + \sum_{i=q+1}^{n}\Bigg[\left(\frac{n}{2}\right)\left(\frac{Z_{i(q+1)j}^2 f'(Z_{i(q+1)j}) + Z_{i(q+1)j}f(Z_{i(q+1)j})}{F(Z_{i(q+1)j})} - \left(\frac{Z_{i(q+1)j}f(Z_{i(q+1)j}}{F(Z_{i(q+1)j})}\right)^2\right) \\
& - \left(\frac{n-2}{2}\right)\left(\frac{Z_{i(q+1)j}^2 f'(Z_{i(q+1)j}) + Z_{i(q+1)j}f(Z_{i(q+1)j})}{1-F(Z_{i(q)j})} + \left(\frac{Z_{i(q+1)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q+1)j})}\right)^2\right) \\
& + \frac{Z_{i(q+1)j}^2 f''(Z_{i(q+1)j}) + Z_{i(q+1)j}f'(Z_{i(q+1)j})}{f(Z_{i(q+1)j})} - \left(\frac{Z_{i(q+1)j}f'(Z_{i(q+1)j}}{f(Z_{i(q+1)j})}\right)^2\Bigg],
\end{aligned}
$$

$$\frac{\partial^2 L_{MRSS2}}{\partial\sigma^2}=\frac{mn}{\sigma^2}$$

$$+\frac{1}{\sigma^2}\sum_{j=1}^{m}\left\{\sum_{i=1}^{q}\left[\left(\frac{n-2}{2}\right)\left(\frac{Z_{i(q)j}^2 f'(Z_{i(q)j})+2Z_{i(q)j}f(Z_{i(q)j})}{F(Z_{i(q)j})}-\left(\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{F(Z_{i(q)j})}\right)^2\right)\right.\right.$$

$$-\frac{n}{2}\left(\frac{Z_{i(q)j}^2 f'(Z_{i(q)j})+2Z_{i(q)j}f(Z_{i(q)j})}{1-F(Z_{i(q)j})}+\left(\frac{Z_{i(q)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q)j})}\right)^2\right)$$

$$\left.+\frac{Z_{i(q)j}^2 f''(Z_{i(q)j})+2Z_{i(q)j}f'(Z_{i(q)j})}{f(Z_{i(q)j})}-\left(\frac{Z_{i(q)j}f'(Z_{i(q+1)j}}{f(Z_{i(q)j})}\right)^2\right]$$

$$+\sum_{i=q+1}^{n}\left[\left(\frac{n}{2}\right)\left(\frac{Z_{i(q+1)j}^2 f'(Z_{i(q+1)j})+2Z_{i(q+1)j}f(Z_{i(q+1)j})}{F(Z_{i(q+1)j})}-\left(\frac{Z_{i(q+1)j}f(Z_{i(q+1)j}}{F(Z_{i(q+1)j})}\right)^2\right)\right.$$

$$-\left(\frac{n-2}{2}\right)\left(\frac{Z_{i(q+1)j}^2 f'(Z_{i(q+1)j})+2Z_{i(q+1)j}f(Z_{i(q+1)j})}{1-F(Z_{i(q)j})}+\left(\frac{Z_{i(q+1)j}f(Z_{i(q+1)j}}{1-F(Z_{i(q+1)j})}\right)^2\right)$$

$$\left.+\frac{Z_{i(q+1)j}^2 f''(Z_{i(q+1)j})+2Z_{i(q+1)j}f'(Z_{i(q+1)j})}{f(Z_{i(q+1)j})}-\left(\frac{Z_{i(q+1)j}f'(Z_{i(q+1)j}}{f(Z_{i(q+1)j})}\right)^2\right].$$

Now, to find the Fisher information, we take the negative expectation and combine like terms to get

$$I_{mn2}(\sigma)=\frac{mn}{2\sigma^2}E\left\{\left[\left(\frac{zf'(z)}{f(z)}\right)^2-\frac{z^2f''(z)+2zf'(z)}{f(z)}\right][g_1(z)+g_2(z)]-2\right\}$$

$$+\frac{mn}{4\sigma^2}E\left\{(n-2)\left[\left(\frac{Z_r^2f'(z)+2zf(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)g_2(z)\right.\right.$$

$$\left.-\left(\frac{z^2f'(z)+2zf(z)}{F(z)}-\left(\frac{zf(z)}{F(z)}\right)^2\right)g_1(z)\right] \tag{3.27}$$

$$+n\left[\left(\frac{z^2f'(z)+2zf(z)}{1-F(z)}+\left(\frac{Z_rf(z)}{1-F(z)}\right)^2\right)g_1(z)\right.$$

$$\left.\left.-\left(\frac{z^2f'(z)+2zf(z)}{F(z)}-\left(\frac{Z_rf(z)}{F(z)}\right)^2\right)g_2(z)\right]\right\}$$

Finally, the asymptotic relative precision of the MRSS estimator of σ with respect to the SRS estimator is

$$\lim_{m\to\infty} RP(\hat{\sigma}_{mlel}, \hat{\sigma}_{ML}) = \frac{I_{mn1}(\sigma)}{I_{mn}(\sigma)} \tag{3.28}$$

3.3. Examples

Example 3.3.1: Let $X_1, X_2, \ldots, X_{mn}$ be random sample of size mn from a normal population with unknown mean, μ and unit variance. We know that the maximum likelihood estimator of μ from this random sample is the sample mean (i.e. $\hat{\mu}_{ML} = \overline{X}$) with Fisher information, $I_{mn}(\mu) = mn$. For a median ranked set sample with odd set size n, we can find the maximum likelihood estimator of μ using equation (3.9) as the solution of the equation

$$\sum_{j=1}^{m}\sum_{i=1}^{n} Z_{i(p)j} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{[2\Phi(Z_{i(p)j})-1]\phi(Z_{i(p)j})}{[1-\Phi(Z_{i(p)j})]\Phi(Z_{i(p)j})} = 0,$$

where $\Phi(.)$ and $\phi(.)$ respectively denote the cumulative distribution function and the probability density function of the standard normal variate.

Using equations (3.10) and (3.11), the Fisher information is

$$\begin{aligned} I_{mn1}(\mu) &= \frac{mn}{\sigma^2}E[g(z)] \\ &+ \frac{mn(n-1)}{2}E\left\{\left(\frac{[\Phi^2(z)+[1-\Phi(z)^2]\phi^2(z)}{[1-\Phi(z)]^2\Phi^2(z)} - \frac{z[2\Phi(z)-1]\phi(z)}{\Phi(z)[1-\Phi(z)]}\right)g(z)\right\}, \end{aligned} \tag{3.29}$$

where $g(z) = \frac{1}{B(p,n-p+1)}\Phi^{p-1}(z)[1-\Phi(z)]^{n-p}$ and $p = (n+1)/2$.

From equation (3.15), the asymptotic relative precision of the MRSS with respect to the simple random sample is

$$\lim_{m\to\infty} RP(\hat{\mu}_{ML_1},\hat{\mu}_{ML}) = E[g(z)] + \frac{(n-1)}{2}E\left\{\left(\frac{[\Phi^2(z)+[1-\Phi(z)^2]\phi^2(z)}{[1-\Phi(z)]^2\Phi^2(z)} - \frac{z[2\Phi(z)-1]\phi(z)}{\Phi(z)[1-\Phi(z)]}\right)g(z)\right\}$$

We estimate the above expectations and all other expectations that follow by numerical integration using Mathematica 2.2 and cross check the results by computer simulation using Matlab 5.2.

For a MRSS with even set size, n, the maximum likelihood estimator, using equation (3.22), is the solution of the equation

$$\sum_{j=1}^{m}\left[\sum_{i=1}^{q} Z_{i(q)j} + \sum_{i=q+1}^{n} Z_{i(q)j}\right] + \frac{1}{2}\sum_{j=1}^{m}\left\{\sum_{i=1}^{q}\frac{[(2n-2)\Phi(Z_{i(q)j})-n+2]\phi(Z_{i(q)j})}{[1-\Phi(Z_{i(q)j})]\Phi(Z_{i(q)j})} + \sum_{i=q+1}^{n}\frac{[(2n-2)\Phi(Z_{i(q+1)j})-n]\phi(Z_{i(q+1)j})}{[1-\Phi(Z_{i(q+1)j})]\Phi(Z_{i(q+1)j})}\right\} = 0,$$

and from equation (3.25), the Fisher information

$$I_{mn2}(\mu) = mnE\{g_1(Z_r)+g_2(Z_r)\} + \frac{mn}{4}E\left\{n\left[\left(\frac{-Z_r\phi(Z_r)}{1-\Phi(Z_r)}+\left(\frac{\phi(Z_r)}{1-\Phi(Z_r)}\right)^2\right)g_1(Z_r) - \left(\frac{-Z_r\phi(Z_r)}{\Phi(Z_r)}+\left(\frac{\phi(Z_r)}{\Phi(Z_r)}\right)^2\right)g_2(Z_r)\right] + (n-2)\left[\left(\frac{-Z_r\phi(Z_r)}{1-\Phi(Z_r)}+\left(\frac{\phi(Z_r)}{1-\Phi(Z_r)}\right)^2\right)g_2(Z_r) - \left(\frac{-Z_r\phi(Z_r)}{\Phi(Z_r)}+\left(\frac{\phi(Z_r)}{\Phi(Z_r)}\right)^2\right)g_1(Z_r)\right]\right\},$$

where $g_1(z) = \dfrac{1}{B(q,q+1)}\Phi^{q-1}(z)[1-\Phi(z)]^q$ and $g_2(z) = \dfrac{1}{B(q+1,q)}\Phi^q(z)[1-\Phi(z)]^{q-1}$ follows from equations (3.23) and (3.24) respectively.

Thus, the asymptotic relative precision

$$\lim_{m\to\infty} RP(\hat{\mu}_{mle2},\hat{\mu}_{ML}) = \frac{I_{mn2}(\mu)}{I_{mn}(\mu)}$$
$$= E\{g_1(z) + g_2(z)\}$$
$$+\frac{1}{4}E\left\{n\left[\left(\left(\frac{-z\phi(z)}{1-\Phi(z)}+\left(\frac{\phi(z)}{1-\Phi(z)}\right)^2\right)g_1(Z_r)-\left(\left(\frac{-z\phi(z)}{\Phi(z)}+\left(\frac{\phi(z)}{\Phi(z)}\right)^2\right)g_2(z)\right]\right.\right.$$
$$\left.+(n-2)\left[\left(\left(\frac{-z\phi(z)}{1-\Phi(z)}+\left(\frac{\phi(z)}{1-\Phi(z)}\right)^2\right)g_2(z)-\left(\left(\frac{-z\phi(z)}{\Phi(z)}+\left(\frac{\phi(z)}{\Phi(z)}\right)^2\right)g_1(z)\right]\right]\right\}$$

We present a comparison of the asymptotic relative precision for RSS (Stokes [45]), that for MRSS, the relative precision for the non-parametric RSS (Dell and Clutter [10]) and the relative precision for the non-parametric MRSS (Muttlak[25]) in Table 3.1.

TABLE 3.1. RELATIVE PRECISION of RSS and MRSS MAXIMUM LIKELIHOOD ESTIMATORS and the NON-PARAMETRIC RSS and MRSS ESTIMATORS FOR μ

Set size n	Asymptotic relative precision (Maximum likelihood)		Relative precision (non-parametric RSS)	
	RSS	MRSS	RSS	MRSS
2	1.48	1.48	1.47	1.47
3	1.96	2.23	1.91	2.23
4	2.44	2.78	2.35	2.77
5	2.92	3.49	2.77	3.49
6	3.40	4.07	3.19	4.06
7	3.88	4.75	3.59	4.75
8	4.36	5.34	4.00	5.34

From this table, it is clear that the maximum likelihood estimator from the MRSS is more precise than its RSS counterpart and of course, is more efficient than the SRS estimator is. However, it is not an improvement upon the non-parametric estimator.

Example 3.3.2: Suppose that we have the same setup of example 3.1 except that the population under consideration is normal, of mean zero and an unknown variance (i.e.

$N(0, \sigma^2)$. Then the SRS maximum likelihood estimator, $\hat{\sigma}_{ML}$ of σ is the sample standard deviation, s and the Fisher information $I_{mn}(\sigma) = 2mn/\sigma^2$.

For a MRSS with odd set size, n, equation (3.16) gives the maximum likelihood estimator, $\hat{\sigma}_{mle1}$ of σ as the solution of the equation

$$mn-(n-1)\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{[2\Phi(Z_{i(p)j})-1]Z_{i(p)j}\phi(Z_{i(p)j})}{[1-\Phi(Z_{i(p)j})]\Phi(Z_{i(p)j})}-\sum_{j=1}^{m}\sum_{i=1}^{n}Z_{i(p)j}^2=0,$$

and the Fisher information follows from equation (3.17) as

$$I_{mn1}(\sigma)=\frac{mn}{\sigma^2}E\left[3z^2g(z)-1\right]+\frac{mn(n-1)}{2\sigma^2}E\left\{\left[\frac{z\phi(z)[2\Phi(z)-1](1-z^2)}{[1-\Phi(z)]\Phi(z)}\right.\right.$$
$$\left.\left.+\frac{[[1-\Phi(z)]^2+\Phi^2(z)]z^2\phi^2(z)}{[1-\Phi(z)]^2\Phi^2(z)}\right]g(z)\right\}.$$

Thus, the asymptotic precision of the MRSS relative to SRS is

$$\lim_{m\to\infty}RP(\hat{\sigma}_{mle1},\hat{\sigma}_{ML})=\frac{1}{2}E\left[3Z_r^2g(Z_r)-1\right]+\frac{(n-1)}{4}E\left\{\left[\frac{Z_r\phi(Z_r)[2\Phi(Z_r)-1](1-Z_r^2)}{[1-\Phi(Z_r)]\Phi(Z_r)}\right.\right.$$
$$\left.\left.+\frac{[[1-\Phi(Z_r)]^2+\Phi^2(Z_r)]Z_r^2\phi^2(Z_r)}{[1-\Phi(Z_r)]^2\Phi^2(Z_r)}\right]g(Z_r)\right\}.$$

However, if the set size is even, then from equation (3.26), the maximum likelihood estimator, $\hat{\sigma}_{mle2}$ of σ is the solution of

$$mn-\frac{1}{2}\sum_{j=1}^{m}\left\{\sum_{i=1}^{q}\frac{Z_{i(q)j}\phi(Z_{i(q)j})[(2n-1)\Phi(Z_{i(q)j})-n+1][2-Z_{i(q)j}^2]}{[1-\Phi(Z_{i(q)j})]\Phi(Z_{i(q)j})}\right.$$
$$\left.+\sum_{i=q+1}^{n}\frac{Z_{i(q+1)j}\phi(Z_{i(q+1)j})[(2n-1)\Phi(Z_{i(q+1)j})-n][2-Z_{i(q+1)j}^2]}{[1-\Phi(Z_{i(q+1)j})]\Phi(Z_{i(q+1)j})}\right\}=0,$$

and the corresponding Fisher information is obtained from equation (3.27) as

$$I_{mn2}(\sigma)=\frac{mn}{2\sigma^2}E\{3Z_r^2[g_1(Z_r)+g_2(Z_r)]-2\}$$

$$+\frac{mn}{4\sigma^2}E\Bigg\{(n-2)\Bigg[\left(\frac{-Z_r^3\phi(Z_r)+2Z_r\phi(Z_r)}{1-\Phi(Z_r)}+\left(\frac{Z_r\phi(Z_r)}{1-\Phi(Z_r)}\right)^2\right)g_2(Z_r)$$

$$-\left(\frac{-Z_r^3\phi(Z_r)+2Z_r\phi(Z_r)}{\Phi(Z_r)}-\left(\frac{Z_r\phi(Z_r)}{\Phi(Z_r)}\right)^2\right)g_1(Z_r)\Bigg] \tag{3.30}$$

$$+n\Bigg[\left(\frac{-Z_r^3\phi(Z_r)+2Z_r\phi(Z_r)}{1-\Phi(Z_r)}+\left(\frac{Z_r\phi(Z_r)}{1-\Phi(Z_r)}\right)^2\right)g_1(Z_r)$$

$$-\left(\frac{-Z_r^3\phi(Z_r)+2Z_r\phi(Z_r)}{\Phi(Z_r)}-\left(\frac{Z_r\phi(Z_r)}{\Phi(Z_r)}\right)^2\right)g_2(Z_r)\Bigg]\Bigg\}.$$

Thus, the asymptotic relative precision is given by

$$\lim_{m\to\infty} RP(\hat{\sigma}_{mle2},\hat{\sigma}_{ML})=\frac{1}{4}E\{3z^2[g_1(z)+g_2(z)]-2\}$$

$$+\frac{1}{8}E\Bigg\{(n-2)\Bigg[\left(\frac{-z^3\phi(z)+2z\phi(z)}{1-\Phi(z)}+\left(\frac{z\phi(z)}{1-\Phi(z)}\right)^2\right)g_2(z)$$

$$-\left(\frac{-z^3\phi(z)+2z\phi(z)}{\Phi(z)}-\left(\frac{z\phi(z)}{\Phi(z)}\right)^2\right)g_1(z)\Bigg]$$

$$+n\Bigg[\left(\frac{-z^3\phi(z)+2z\phi(z)}{1-\Phi(z)}+\left(\frac{z\phi(z)}{1-\Phi(z)}\right)^2\right)g_1(z)$$

$$-\left(\frac{-z^3\phi(z)+2z\phi(z)}{\Phi(z)}-\left(\frac{z\phi(z)}{\Phi(z)}\right)^2\right)g_2(z)\Bigg]\Bigg\}.$$

Table 3.2 compares the asymptotic relative precision for maximum likelihood estimation under RSS (Stokes [45]), the asymptotic relative precision under the non-parametric RSS (Stokes [42]) and that for maximum likelihood estimation under MRSS. The results show that for the odd set sizes, the maximum likelihood estimator from MRSS is worse than the maximum likelihood estimator from a simple random sample. This is in line with the result in Sinha et al. [40], where they showed that the use of the traditional median in the estimation of the normal variance results is a poor estimator as compared to the simple random sample estimator. For even set sizes however, the MRSS estimator is a little more efficient that that of the SRS (note the departure from the traditional median in this case). On the other hand the RSS estimator is more efficient than that of the MRSS over all set sizes except $n = 2$, when they are equal.

TABLE 3.2. RELATIVE PRECISION OF MAXIMUM LIKELIHOOD ESTIMATORS OF σ FROM RSS, MRSS AND THE NON-PARAMETRIC RSS FOR $N(0, \sigma^2)$

Set size	Asymptotic relative precision		
n	mle under RSS	mle under MRSS	Non-parametric RSS
2	1.14	1.14	1.00
3	1.27	0.98	1.08
4	1.41	1.08	1.18
5	1.54	0.98	1.27
6	1.68	1.05	1.38
7	1.81	0.99	1.48
8	1.95	1.04	1.57

Example 3.3.3: Let $F(x) = 1 - \exp[-x/\sigma]$ and $f(x) = (1/\sigma)\exp[-x/\sigma]$ be the cdf and pdf of the exponential distribution, where $0 \le x < \infty$. Then for odd set size, we use equation (3.16) to obtain the maximum likelihood estimator of σ, $\hat{\sigma}_{mle1}$ as the solution of the equation

$$mn\left[1+\frac{mn(n-1)}{2}\right]+\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{\exp[-Z_{i(p)j}]}{1-\exp[-Z_{i(p)j}]}=0.$$

The Fisher information follows from equation (3.17) as

$$I_{mn1}(\sigma)=\frac{mn}{\sigma^2}E[2zg(z)-1]$$
$$+\frac{mn(n-1)}{2\sigma^2}E\left\{\left[\frac{z(2-z)(1-2\exp[-z])}{1-\exp[-z]}+\frac{z^2[(1-\exp(-z))^2+\exp(-2z)]}{(1-\exp[-z])^2}\right]g(z)\right\}, \tag{3.31}$$

where $g(z)=\frac{1}{B(p,n-p+1)}[1-\exp(-z)]^{p-1}\exp[-z(n-p)]$ following from equation (3.11).

From a simple random sample of size mn, the maximum likelihood estimator of σ is the sample mean, $\overline{X}_{SRS}$ and the Fisher information $I_{mn}(\sigma) = mn/\sigma^2$. Hence, the asymptotic relative precision is

$$\lim_{m\to\infty} RP(\hat{\sigma}_{mle1}, \overline{X}_{SRS}) = E[2zg(z)-1] + \frac{(n-1)}{2} E\left\{\left[\frac{z(2-z)(1-2\exp[-z])}{1-\exp[-z]} + \frac{z^2[(1-\exp(-z))^2+\exp(-2z)]}{(1-\exp[-z])^2}\right] g(z)\right\}.$$

The expectations here are also found by numerical integration and cross-checked by computer simulation as in the previous example.

Similarly for an even set size, we can obtain the maximum likelihood estimator by substituting for F and f in equation (3.26) and solving for σ by iteration. The Fisher information, using equation (3.27) is given by

$$\begin{aligned} I_{mn2}(\sigma) = {} & \frac{mn}{2\sigma^2} E\{2Z_r[g_1(Z_r)+g_2(Z_r)]-2\} \\ & + \frac{mn}{4\sigma^2} E\left\{(n-2)\left[2Z_r g_2(Z_r) - \left(\frac{(-Z_r^2+2Z_r)\exp(-Z_r)-2Z_r\exp(-2Z_r)}{[1-\exp(-Z_r)]^2}\right) g_1(Z_r)\right]\right. \\ & \left. n\left[2Z_r g_1(Z_r) - \left(\frac{(-Z_r^2+2Z_r)\exp(-Z_r)-2Z_r\exp(-2Z_r)}{[1-\exp(-Z_r)]^2}\right) g_2(Z_r)\right]\right\}. \end{aligned} \tag{3.32}$$

and the asymptotic relative precision or the MRSS estimator relative to the SRS estimator is

$$\begin{aligned} \lim_{m\to\infty} RP(\hat{\sigma}_{mle2}, \overline{X}) = {} & \frac{1}{2} E\{2Z_r[g_1(Z_r)+g_2(Z_r)]-2\} \\ & + \frac{1}{4} E\left\{(n-2)\left[2Z_r g_2(Z_r) - \left(\frac{(-Z_r^2+2Z_r)\exp(-Z_r)-2Z_r\exp(-2Z_r)}{[1-\exp(-Z_r)]^2}\right) g_1(Z_r)\right]\right. \\ & \left. +n\left[2Z_r g_1(Z_r) - \left(\frac{(-Z_r^2+2Z_r)\exp(-Z_r)-2Z_r\exp(-2Z_r)}{[1-\exp(-Z_r)]^2}\right) g_2(Z_r)\right]\right\}. \end{aligned}$$

Table 3.3 compares the relative efficiency values for the non-parametric RSS (Dell and Clutter [10]) and non-parametric MRSS (Muttlak [25]), the values for maximum likelihood estimation under RSS (Stokes [45]) and those for maximum likelihood estimation under MRSS.

TABLE 3.3. RELATIVE PRECISION VALUES for Exp(σ)

Set size n	Asymptotic relative precision (Maximum likelihood)		Relative precision Non-parametric methods	
	RSS	MRSS	RSS	MRSS
2	1.40	1.40	1.33	1.33
3	1.81	1.92	1.64	2.25
4	2.21	2.37	1.92	2.44
5	2.62	2.88	2.19	2.23
6	3.02	3.33	2.45	2.14
7	3.42	3.83	2.70	1.80
8	3.83	4.29	2.94	1.67

It is clear here that the MRSS maximum likelihood estimator dominates all the other estimators except at n = 3 and n = 4, when the non-parametric MRSS estimator does better.

Example 3.3.4: In this example, we consider a more general form of example 3.3.3. Let the gamma distribution function be of the form $F(x)=\frac{1}{\Gamma(\alpha)}\int_0^{x/\sigma} t^{\alpha-1}e^{-t}, \alpha>0$. The probability density function of this distribution is $f(x)=\frac{1}{\sigma\Gamma(\alpha)}\left(\frac{x}{\sigma}\right)^{\alpha-1}\exp(-x/\sigma)$. This clearly yields the distribution of example 3.3.3 when $\alpha=1$. We will find the asymptotic relative precision for MRSS and RSS and compare them for some specific values of α.

The maximum likelihood estimators of σ for odd and even set-sized MRSS can respectively be found by suitably substituting for *F(z)* and *f(z)* into equations (3.16) and

(3.26), and then solving for σ by iteration. The simple random sample Fisher information for σ is given by

$$I_{mn}(\sigma)=\frac{\alpha mn}{\sigma^2}. \tag{3.33}$$

Note that for this family of distributions,

$$f'(z)=\left(\frac{\alpha-1}{z}-1\right)f(z) \text{ and } f''(z)=\left[-\frac{\alpha-1}{z^2}+\left(\frac{\alpha-1}{z}-1\right)^2\right]f(z).$$

This implies

$$\frac{z^2f''(z)+2zf'(z)}{f(z)}-\left(\frac{zf'(z)}{f(z)}\right)^2=\alpha-1-2z \text{ and } z^2f'(z)+2zf(z)=[(\alpha+1)z-z^2]f(z).$$

For odd and even set-sized MRSS, the respective Fisher information are given by equations (3.17) and (3.27) as

$$\begin{aligned}I_{mn1}(\sigma)=\frac{mn}{\sigma^2}E\{[\alpha-1-2z]g(z)-1\}+\frac{mn(n-1)}{2\sigma^2}E\Bigg\{\Bigg[&\frac{[2F(z)-1][(\alpha+1)z-z^2]f(z)}{F(z)[1-F(z)]}\\&+\frac{[F^2(z)+(1-F(z))^2]z^2f^2(z)}{F^2(z)[1-F(z)]^2}\Bigg]g(z)\Bigg\}\end{aligned} \tag{3.34}$$

and

$$\begin{aligned}I_{mn2}(\sigma)=&\frac{mn}{2\sigma^2}E\{\alpha-1-2z](g_1(z)+g_2(z))-2\}\\&+\frac{mn}{4\sigma^2}E\Bigg\{(n-2)\Bigg[\left(\frac{[(\alpha+1)z-z^2]f(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)g_2(z)-\left(\frac{[(\alpha+1)z-z^2]f(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right)\Bigg]g_1(z)\\&+n\Bigg[\left(\frac{[(\alpha+1)z-z^2]f(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)g_1(z)-\left(\frac{[(\alpha+1)z-z^2]f(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right)g_2(z)\Bigg]\Bigg\}.\end{aligned} \tag{3.34}$$

Consequently, the corresponding asymptotic relative precision are give by

$$\lim_{m\to\infty} RP(\hat{\sigma}_{mle1},\hat{\sigma}_{ML}) = \frac{1}{\alpha}E\{[\alpha-1-2z]g(z)-1\}$$

$$+\frac{(n-1)}{2\alpha}E\left\{\left[\frac{[2F(z)-1][(\alpha+1)z-z^2]f(z)}{F(z)[1-F(z)]}+\frac{[F^2(z)+(1-F(z))^2]z^2f^2(z)}{F^2(z)[1-F(z)]^2}\right]g(z)\right\},$$

and

$$\lim_{m\to\infty} RP(\hat{\sigma}_{mle2},\hat{\sigma}_{ML}) = \frac{1}{2\alpha}E\{\alpha-1-2z](g_1(z)+g_2(z))-2\}$$

$$+\frac{1}{4\alpha}E\left\{(n-2)\left[\left(\frac{[(\alpha+1)z-z^2]f(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)g_2(z)-\left(\frac{[(\alpha+1)z-z^2]f(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right)\right]g_1(z)\right.$$

$$\left.+n\left[\left(\frac{[(\alpha+1)z-z^2]f(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right)g_1(z)-\left(\frac{[(\alpha+1)z-z^2]f(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right)g_2(z)\right]\right\}.$$

We present the asymptotic relative precision for $\alpha = 2$ and $\alpha = 3$ in tables 3.4 and 3.5 respectively for both MRSS and RSS.

From Stokes [45], it is easy to show that the Fisher information for σ from RSS is

$$I^{*}_{mn}(\sigma) = \frac{\alpha mn}{\sigma^2}+\frac{mn(n-1)}{\sigma^2}E\left\{\frac{[zf(z)]^2}{F(z)(1-F(z))}\right\} \tag{3.35}$$

Thus, the asymptotic relative precision for RSS is

$$\lim_{m\to\infty} RP(\hat{\sigma}^{*}_{ML},\hat{\sigma}_{ML}) = 1+\frac{(n-1)}{\alpha}E\left\{\frac{[zf(z)]^2}{F(z)(1-F(z))}\right\}$$

TABLE 3.4. COMPARISON of RELATIVE PRECISION VALUES for the ESTIMATION of σ from GAMMA (2.0)

Set size n	Asymptotic relative precision (maximum likelihood)		Relative precision (non-parametric methods)	
	RSS	MRSS	RSS	MRSS
2	1.44	1.44	1.39	1.39
3	1.88	2.07	1.75	2.23
4	2.32	2.56	2.10	2.56
5	2.76	3.16	2.42	2.64
6	3.20	3.67	2.74	2.70
7	3.64	4.26	3.05	2.48
8	4.07	4.78	3.35	2.40

All the required expectations were found as in the previous examples. Table 3.4 compares for gamma (2.0), the asymptotic relative precision for maximum likelihood estimation of σ under RSS and MRSS, and the relative precision under the non-parametric RSS and MRSS. Table 3.5 shows the same comparison for gamma (3.0). The relative precision for the non-parametric RSS and MRSS were obtained from Dell and Clutter [10] and Muttlak [25], computations being made by us for the extra set sizes not reported in those papers.

TABLE 3.5. COMPARISON OF RELATIVE PRECISION VALUES FOR THE ESTIMATION OF σ FROM GAMMA (3.0)

Set size n	Asymptotic relative precision (maximum likelihood)		Relative precision (non-parametric methods)	
	RSS	MRSS	RSS	MRSS
2	1.45	1.45	1.41	1.41
3	1.90	2.12	1.80	2.24
4	2.36	2.63	2.16	2.62
5	2.81	3.26	2.52	2.85
6	3.26	3.80	2.87	3.01
7	3.71	4.42	3.20	2.91
8	4.17	4.96	3.54	2.90

The trend in these results is the same as that observed under the exponential distribution. That is besides at the set size of three, when the non-parametric MRSS estimator dominates, the MRSS maximum likelihood estimator is the overall dominant.

3.4. Two-parameter family

We will now briefly discuss the case where neither of the location and scale parameters is known, i.e. the two-parameter family. In this case, the usual principle of maximum likelihood requires us to obtain the first derivative of the appropriate loglikelihood function with respect to each parameter, set each result equal to zero, and simultaneously solve the resulting equations for the parameters. For instance, if μ and σ are not known and we have a MRSS of odd set size, we obtain their maximum likelihood estimates by simultaneously solving equation (3.9) and equation (3.16). Similarly, if we have an even set size, we set equation (3.22) equal to zero and solve simultaneously with equation (3.26) as in the case of the odd set size.

We will now investigate the performance of these estimators for the case of odd set sizes against the performance of the corresponding SRS estimators. We will do this using the Fisher information matrix, which has $I_{mn1}(\mu)$ and $I_{mn1}(\sigma)$ (equations (3.10) and (3.17) respectively) as its diagonal elements. The off diagonal elements are given by

$$-E\left\{\frac{\partial^2 L_{MRSS1}}{\partial\mu\,\partial\sigma}\right\}=\frac{mn}{\sigma^2}E\left\{\left[z\left(\frac{f'(z)}{f(z)}\right)^2-\frac{zf''(z)+f'(z)}{f(z)}\right]g(z)\right\}$$
$$+\frac{mn(n-1)}{2\sigma^2}E\left\{\left[\frac{[2F(z)-1][zf'(z)+f(z)]}{F(z)[1-F(z)]}+\frac{z[F^2(z)+(1-F(z))^2]f^2(z)}{F^2(z)[1-F(z)]^2}\right]g(z)\right\}$$

where $\frac{1}{B(p,p)}F^{p-1}(z)[1-F(z)]^{n-p}$ and $p=\frac{n+1}{2}$. Just as in the case of RSS (Stokes [49]), the off-diagonal elements are zero for symmetric distributions. Note that $\frac{\partial^2 L_{MRSS1}}{\partial\mu\,\partial\sigma}=\frac{\partial^2 L_{MRSS1}}{\partial\sigma\,\partial\mu}$.

In the case of SRS, the diagonal elements are given by $I_{mn}(\mu)$ and $I_{mn}(\sigma)$, the Fisher information for μ and σ respectively from a simple random sample. The off-diagonal elements are given by

$$-E\left\{\frac{\partial^2 L}{\partial\mu\,\partial\sigma}\right\}=E\left\{\frac{zf''(z)+f'(z)}{f(z)}-z\left(\frac{f'(z)}{f(z)}\right)^2\right\}$$

This is also zero for symmetric distributions.

Thus, to compare the MRSS estimators with those of the SRS, we compare the determinants of the information matrices

$$\begin{bmatrix} I_{mn1}(\mu) & -E\left\{\dfrac{\partial^2 L_{MRSS1}}{\partial\mu\,\partial\sigma}\right\} \\ -E\left\{\dfrac{\partial^2 L_{MRSS1}}{\partial\mu\,\partial\sigma}\right\} & I_{mn1}(\sigma) \end{bmatrix} \text{ and } \begin{bmatrix} I_{mn}(\mu) & -E\left\{\dfrac{\partial^2 L}{\partial\mu\,\partial\sigma}\right\} \\ -E\left\{\dfrac{\partial^2 L}{\partial\mu\,\partial\sigma}\right\} & I_{mn}(\sigma) \end{bmatrix},$$

which are the information matrices for the MRSS and RSS estimators respectively. This clearly shows that the trend observed in the relative precision values in the case of the one-parameter family also holds here even though the actual estimates may differ. This observation agrees with Stokes [45] in the case of RSS. The analyses for even set sizes follow similarly.

CHAPTER 4

SOME UNBIASED ESTIMATORS OF THE LOCATION-SCALE PARAMETERS USING MEDIAN RANKED SET SAMPLING

In this chapter, we propose some unbiased estimators of the location scale parameters of the family of distributions considered in the previous chapter in terms of the median ranked set samples. We do this following the definitions of Lloyd [18] and then modifying the best linear unbiased estimators (BLUE's) presented by Stokes [45]. We compare our proposed estimators with the maximum likelihood estimators under MRSS (Chapter 3) and those under RSS (Stokes [45]). Recall that in using the method of MRSS, the order statistics of interested in $(n+1/2)^{th}$, $(n/2)^{th}$ and the $((n+2)/2)^{th}$ order statistics from a set of size n are the, the last two in case of even set sizes and the first one in case of odd ones.

4.1. Notation and some useful results

For simplicity, let the functions $g(z)$, $g_1(z)$ and $g_2(z)$ be defined as in equations (3.11), (3.23) and (3.24) respectively. Further, let

$$E_1(n)=E\left\{\left[\left(\frac{f'(z)}{f(z)}\right)^2-\frac{f''(z)}{f(z)}\right]g(z)\right\},$$

$$E_2(n)=E\left[\left\{\frac{[2F(Z_r)-1]f'(Z_r)}{F(Z_r)[1-F(Z_r)]}+\frac{[F^2(Z_r)+(1-F(Z_r))^2]f^2(Z_r)}{[1-F(Z_r)]^2F^2(Z_r)}\right\}g(Z_r)\right],$$

$$\mathrm{E}_3(\mathrm{n})=E\left\{\left[\left(\frac{Z_rf'(Z_r)}{f(Z_r)}\right)^2-\frac{Z_r^2f''(Z_r)+2Z_rf'(Z_r)}{f(Z_r)}\right]g(z)-1\right\},$$

$$E_4(n)=E\left[\left\{\frac{[2F(Z_r)-1][Z_r^2f'(Z_r)+2Z_rf(Z_r)]}{[1-F(Z_r)]F(Z_r)}+\frac{[F^2(Z_r)+(1-F(Z_r))^2]Z_{i(p)j}^2f^2(Z_r)}{F^2(Z_r)[1-F(Z_r)]^2}\right\}g(z)\right].$$

Also let

$$E_1'(n)=E\left\{\left[\left(\frac{f'(z)}{f(z)}\right)^2-\frac{f''(z)}{f(z)}\right][g_1(z)+g_2(z)]\right\},$$

$$E_2'=E\left\{n\left[\left(\frac{f'(Z_r)}{1-F(Z_r)}+\left(\frac{f(Z_r)}{1-F(Z_r)}\right)^2\right)g_1(Z_r)-\left(\frac{f'(Z_r)}{F(Z_r)}-\left(\frac{f(Z_r)}{F(Z_r)}\right)^2\right)g_2(Z_r)\right]\right.$$
$$\left.+(n-2)\left[\left(\frac{f'(Z_r)}{1-F(Z_r)}+\left(\frac{f(Z_r)}{1-F(Z_r)}\right)^2\right)g_2(Z_r)-\left(\frac{f'(Z_r)}{F(Z_r)}-\left(\frac{f(Z_r)}{F(Z_r)}\right)^2\right)g_1(Z_r)\right]\right\},$$

$$E_3'=E\left\{\left[\left(\frac{Z_rf'(Z_r}{f(Z_r)}\right)^2-\frac{Z_r^2f''(Z_r)+2Z_rf'(Z_r)}{f(Z_r)}\right][g_1(Z_r)+g_2(Z_r)]-2\right\},$$

and

$$E_4'(n)=E\left\{(n-2)\left[\left[\left(\frac{Z_r^2f'(Z_r)+2Z_rf(Z_r)}{1-F(Z_r)}+\left(\frac{Z_rf(Z_r}{1-F(Z_r)}\right)^2\right)g_2(Z_r)-\left(\frac{Z_r^2f'(Z_r)+2Z_rf(Z_r)}{F(Z_r)}-\left(\frac{Z_rf(Z_r}{F(Z_r)}\right)^2\right)g_1(Z_r)\right]\right.\right.$$

$$\left.\left.+n\left[\left(\frac{Z_r^2f'(Z_r)+2Z_rf(Z_r)}{1-F(Z_r)}+\left(\frac{Z_rf(Z_r}{1-F(Z_r)}\right)^2\right)g_1(Z_r)-\left(\frac{Z_r^2f'(Z_r)+2Z_rf(Z_r)}{F(Z_r)}-\left(\frac{Z_rf(Z_r}{F(Z_r)}\right)^2\right)g_2(Z_r)\right]\right\}.$$

Then the information relations of equations (3.10), (3.17), (3.25) and (3.26) can respectively be written as

$$I_{mn1}(\mu)=\frac{mn}{\sigma_2}E_1(n)+\frac{mn(n-1)}{2\sigma^2}E_2(n), \tag{4.1}$$

$$I_{mn1}(\sigma)=\frac{mn}{\sigma_2}E_3(n)+\frac{mn(n-1)}{2\sigma^2}E_4(n), \tag{4.2}$$

$$I_{mn2}(\mu)=\frac{mn}{2\sigma_2}E_1'(n)+\frac{mn}{4\sigma^2}E_2'(n), \tag{4.3}$$

$$I_{mn2}(\sigma)=\frac{mn}{2\sigma_2}E_3'(n)+\frac{mn}{4\sigma^2}E_4'(n). \tag{4.4}$$

We numerically compute and present the values of the E_i's and the E_i''s in Tables 4.1 and 4.2 for different sample sizes and different distributions for use in the examples in this chapter. The numerical computation is done using Mathematica 2.2.

4.2. The Unbiased Estimators and their Performance

Following Lloyd [18] as in Stokes [45], we define $\alpha_{(.:n)}=E[Z_{i(.)}]$ and $v_{(.:n)}=Var[Z_{i(.)}]$, where $Z_{i(.)}=\dfrac{X_{i(.)}-\mu}{\sigma}$ and $X_{i(.)}$ is the $(.)^{th}$ order statistic in the i^{th} set. Thus, it follows that $E[X_{i(.:n)}]=\mu+\sigma\alpha_{(.:n)}$ and $Var[X_{i(.:n)}]=\sigma^2v_{(.:n)}$. We will maintain as in the previous chapter that $p=(n+1)/2$ and $q=n/2$.

Now suppose that we have an odd set size, n, and that σ is known. Then the proposed estimator of μ from a MRSS with m cycles is

$$\hat{\mu}_{UB_1} = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} X_{i(p)} - \sigma\alpha_{(p:n)} . \tag{4.5}$$

This proposed estimator:

(i) is unbiased for μ and

(ii) has variance

$$Var[\hat{\mu}_{UB_1}] = \frac{\sigma^2}{mn} v_{(p:n)} . \tag{4.6}$$

To show (i), we take the expected value of $\hat{\mu}_{UB_1}$ to get

$$E[\hat{\mu}_{UB_1}] = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} E[X_{i(p)j}] - \sigma\alpha_{(p:n)} = \frac{1}{mn} mn[\mu + \alpha_{(p,n)}\sigma] - \alpha_{(p,n)}\sigma = \mu .$$

To show (ii), we take the variance of of $\hat{\mu}_{UB_1}$ as follows

$$\begin{aligned}
Var[\hat{\mu}_{UB_1}] = Var\left[\frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} X_{i(p)j} - \sigma\alpha_{(p:n)}\right] &= \frac{1}{m^2n^2}\sum_{j=1}^{m}\sum_{i=1}^{n} Var[X_{i(p)j}] \\
&= \frac{1}{m^2n^2}\sum_{j=1}^{m}\sum_{i=1}^{n} Var[\mu + \sigma Z_{i(p)j}] \\
&= \frac{1}{m^2n^2}\sum_{j=1}^{m}\sum_{i=1}^{n} \sigma^2 Var[Z_{i(p)j}] \\
&= \frac{1}{m^2n^2} mn\sigma^2 v_{(p:n)} \\
&= \frac{\sigma^2 v_{(p:n)}}{mn}.
\end{aligned}$$

The relative efficiency of this estimator with respect to the maximum likelihood estimator is

$$eff(\hat{\mu}_{UB_1},\hat{\mu}_{ML_1})=\frac{1}{I_{mn1}(\hat{\mu}_{ML_1})Var[\hat{\mu}_{UB_1}]}$$
$$=\frac{1}{[E_1(n)+0.5(n-1)E_2(n)]v_{(p:n)}}. \tag{4.7}$$

However, if we have an even set size, then the unbiased estimator is

$$\hat{\mu}_{UB_2}=\frac{1}{mn}\sum_{j=1}^{m}\left[\sum_{i=1}^{q}X_{i(q)}+\sum_{i=q+1}^{n}X_{i(q+1)}\right]-\frac{\sigma}{2}\left(\alpha_{(q:n)}+\alpha_{(q+1:n)}\right), \tag{4.8}$$

with variance

$$Var[\hat{\mu}_{UB_2}]=\frac{\sigma^2}{2mn}\left(v_{(q:n)}+v_{(q+1:n)}\right). \tag{4.9}$$

The unbiasedness of this estimator and its variance can easily be shown in a manner similar to the case of the odd set size above. The corresponding relative efficiency is

$$eff(\hat{\mu}_{UB_2},\hat{\mu}_{ML_2})=\frac{2}{[0.5E_1'(n)+0.25E_2'(n)](v_{(q:n)}+v_{(q+1:n)})}. \tag{4.10}$$

For symmetric distributions,

$$Var[\hat{\mu}_{UB_2}]=\frac{\sigma^2}{mn}v_{(q:n)}=\frac{\sigma^2}{mn}v_{(q+1:n)}, \tag{4.11}$$

and then the relative efficiency is given by

$$eff(\hat{\mu}_{UB_2},\hat{\mu}_{ML_2})=\frac{1}{[0.5E_1'(n)+0.25E_2'(n)](v_{(q:n)})}=\frac{1}{[0.5E_1'(n)+0.25E_2'(n)](v_{(q+1:n)})} \tag{4.12}$$

Suppose now that μ is known and we wish to estimate σ using MRSS. The proposed estimator for σ is

$$\hat{\sigma}_{UB_1}=\frac{1}{\alpha_{(p:n)}}\left\{\frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n}X_{i(p)j}-\mu\right\}, \tag{4.13}$$

This proposed estimator, $\hat{\sigma}_{UB_1}$:

(i) is unbiased for σ and

(ii) has variance

$$Var[\hat{\sigma}_{UB_1}] = \frac{\sigma^2 v_{(p:n)}}{mn\alpha^2_{(p:n)}}. \tag{4.14}$$

Consequently, the relative efficiency of this estimator with respect to the MRSS maximum likelihood estimator is

$$eff(\hat{\sigma}_{UB_1}, \hat{\sigma}_{ML_1}) = \frac{\alpha^2_{(p:n)}}{[0.5E_3(n)+0.25(n-1)E_4(n)]v_{(p:n)}}. \tag{4.15}$$

It is clear that the results of equation (4.13) and (4.14) cannot be used if the underlying distribution is symmetric, as in this case, $\alpha_{(p:n)} = 0$ for all odd set sizes.

For an even set size,

$$\hat{\sigma}_{UB_2} = \frac{1}{mn}\sum_{j=1}^{m}\left[\frac{1}{\alpha_{(q:n)}}\sum_{i=1}^{q} X_{i(q)} + \frac{1}{\alpha_{(q+1:n)}}\sum_{i=q+1}^{n} X_{i(q+1)}\right] - \frac{\mu}{2}\left[\frac{1}{\alpha_{(q:n)}} + \frac{1}{\alpha_{(q+1:n)}}\right] \tag{4.16}$$

is unbiased for σ and has variance

$$Var[\hat{\sigma}_{UB_2}] = \frac{\sigma^2}{2mn}\left[\frac{v_{(q:n)}}{\alpha^2_{(q:n)}} + \frac{v_{(q+1:n)}}{\alpha^2_{(q+1:n)}}\right]. \tag{4.17}$$

Thus, the relative efficiency with respect to the corresponding maximum likelihood estimator is

$$eff(\hat{\sigma}_{UB_2}, \hat{\sigma}_{ML_2}) = \frac{2[\alpha_{(q:n)}\alpha_{(q+1:n)}]^2}{[0.5E_3'(n)+0.25E_4'(n)][\alpha^2_{(q:n)}v_{(q+1:n)} + \alpha^2_{(q+1:n)}v_{(q:n)}]}. \tag{4.18}$$

In case the underlying distribution is symmetric, then

$$Var[\hat{\sigma}_{UB_2}] = \frac{\sigma^2}{mn}\frac{v_{(q:n)}}{\alpha^2_{(q:n)}} = \frac{\sigma^2}{mn}\frac{v_{(q+1:n)}}{\alpha^2_{(q+1:n)}}, \tag{4.19}$$

and

$$
\begin{aligned}
eff(\hat{\sigma}_{UB_2},\hat{\sigma}_{ML_2}) &= \frac{[\alpha_{(q:n)}]^4}{[0.5E_3'(n)+0.25E_4'(n)][\alpha^2_{(q:n)}v_{(q:n)}]} \\
&= \frac{[\alpha_{(q+1:n)}]^4}{[0.5E_3'(n)+0.25E_4'(n)][\alpha^2_{(q+1:n)}v_{(q+1:n)}]}.
\end{aligned}
\tag{4.20}
$$

TABLE 4.1. THE EXPECTATIONS in the INFORMATION RELATIONS for the SCALE PARAMETERS

Set size	2		3		4		5	
Expectation	E_3'	E_4'	E_3	E_4	E_3'	E_4'	E_3	E_4
Normal	4.0000	1.0803	0.3460	1.6204	0.6920	7.2168	-0.1395	1.0543
Exponential	2.0000	1.6165	0.6667	1.2580	1.3333	6.8098	0.5667	1.1542
Gamma(2.0)	4.0000	3.5140	1.6482	2.4840	3.2963	13.8866	1.5436	2.3873
Gamma(3.0)	6.0000	5.4280	2.6425	3.7117	5.2850	20.9922	2.5367	3.6272
Set size	**6**		**7**		**8**			
Expectation	E_3'	E_4'	E_3	E_4	E_3'	E_4'		
Normal	-0.2790	8.9787	-0.3687	0.7807	-0.7373	9.7991		
Exponential	1.1333	11.0633	0.5190	1.1040	1.0381	15.1032		
Gamma(2.0)	3.0872	23.2061	1.4940	2.3413	2.9881	32.2824		
Gamma(3.0)	5.0734	35.4264	2.4866	3.5875	4.9732	49.5947		

TABLE 4.2. The EXPECTATIONS in the INFORMATION RELATIONS FOR THE NORMAL MEAN

Set size	2		3		4		5	
Expectation	E_1'	E_2'	E_1	E_2	E_1'	E_2'	E_1	E_2
Normal	2.0000	1.9923	1.0000	1.2292	1.9996	7.1207	1.0000	1.2435
Set size	**6**		**7**		**8**			
Expectation	E_1'	E_2'	E_1	E_2	E_1'	E_2'		
Normal	2.0018	12.2723	1.0000	1.2508	2.0000	17.3790		

4.3. Examples

Now, we will use the examples of Chapter 3 and compare the estimators obtained therein to the estimators of the present chapter.

Example 4.3.1: Suppose that we have a median ranked set sample from a normal distribution. Then $F = \Phi$, the cumulative distribution function of the standard normal distribution. Suppose further that σ is known and μ is the parameter of interest. Then for an odd set size, n, we estimate μ using equation (4.5). From equation (4.6), we obtain the variance of this estimator. Thus, the efficiency of this estimator relative to the maximum likelihood estimator, $\hat{\mu}_{ML_1}$, is given by

$$eff(\hat{\mu}_{UB_1}, \hat{\mu}_{ML_1}) = \frac{1}{[E_1(n) + 0.5(n-1)E_2(n)]v_{(p:n)}}.$$

For example, if $n = 3$, then $p = 2$ and $v_{(2:3)} = 0.44867$, from Harter and Balakrishnan [12]. From Table 4.2, $E_1 = 1.0$ and $E_2 = 1.2292$.

Hence,

$$eff(\hat{\mu}_{UB_1}, \hat{\mu}_{ML_1}) = \frac{1}{[1+0.5(2)(1.2292)](0.44867)} = 0.9998.$$

Similarly, if n is even and μ is of interest while σ is known, we estimate μ using equation (4.8) and obtain the variance from equation (4.9). The efficiency of this estimator relative to the maximum likelihood estimator is given by equation (4.12).

The efficiency values are shown in Table 4.3 on page 60. We observe from this table that for all of the set sizes, the proposed unbiased estimators are nearly as efficient as the maximum likelihood estimators obtained in Chapter 3 are.

Example 4.3.2: Let $F = \Phi$ and suppose now that the known parameter is μ and the parameter of interest is σ. Then since the underlying distribution is symmetric, we cannot use the proposed estimator for odd set sizes (see equation 4.12). However, we can use equation (4.13), the proposed estimator for even set sizes. Equation (4.16) and equation (4.17) respectively yield the corresponding variance of the estimator and its efficiency relative to the MRSS maximum likelihood estimator. The results are shown in Table 4.3. It is clear from the efficiency values that the proposed estimators are doing very poorly, which implies that they will even do much worse than the simple random sample estimator.

Example 4.3.3: Let $F(x)=1-\exp(-x/\sigma)$ be the cdf of the exponential random variable. Then for odd set size, n, the unbiased estimator of σ and its variance are given by equation (4.13) and equation (4.14) respectively. The efficiency of this estimator relative to its maximum likelihood counterpart is given in equation (4.15). Similarly, for even set sizes, we estimate σ using equation (4.16). We find the variance of this estimator and the corresponding efficiency from equations (4.17) and (4.18) respectively. The relative efficiency values of the proposed estimator with respect to the MLE from Chapter 3 are reported in Table 4.3. The efficiency values indicate that the use of the proposed estimators rather than their maximum likelihood counterparts results in some loss, which is negligible except for a set size of two.

Example 4.3.4: Finally, we consider the gamma distributions of example 3-3.4. Recall that the distribution function is of the form $F(x)=\frac{1}{\Gamma(\alpha)}\int_0^{x/\sigma} t^{\alpha-1}e^{-t}, \alpha>0$ and the probability density function of this distribution is $f(x)=\frac{1}{\sigma\Gamma(\alpha)}\left(\frac{x}{\sigma}\right)^{\alpha-1}\exp(-x/\sigma)$. For any given α, we trace the steps of example 4-3.3 above. The efficiencies of the estimators under gamma (2) and gamma (3) distributions are also shown in Table 4.3. The trend in the performance of the proposed estimators is similar to that observed under the exponential distribution.

TABLE 4.3. EFFICIENCIES of the PROPOSED UNBIASED ESTIMATORS RELATIVE to the MAXIMUM LIKELIHOOD ESTIMATORS under MRSS

		set size						
Distribution		2	3	4	5	6	7	8
Normal(μ, 1)	$eff(\hat{\mu}_{UB}, \hat{\mu}_{ML})$	0.97922	0.99982	0.99793	0.99982	0.99818	0.99986	0.99952
Normal(0, σ)	$eff(\hat{\sigma}_{UB}, \hat{\sigma}_{ML})$	0.20569		0.11404		0.07837		0.05971
Exp (σ)	$eff(\hat{\sigma}_{UB}, \hat{\sigma}_{ML})$	0.91567	0.99915	0.96906	0.99911	0.98377	0.99927	0.98993
Gamma(2.0)	$eff(\hat{\sigma}_{UB}, \hat{\sigma}_{ML})$	0.95276	0.99966	0.98375	0.99974	0.99180	0.99980	0.99504
Gamma(3.0)	$eff(\hat{\sigma}_{UB}, \hat{\sigma}_{ML})$	0.96534	0.99974	0.98848	0.99979	0.99430	0.99985	0.99660

Note: The numerical subscripts of the proposed unbiased estimators have been dropped for convenience. We however assume the numerical subscript 1 for odd set sizes and 2 for even set sizes.

4.4. Comparison with the MLE under RSS

We now wish to compare the proposed unbiased estimators of the previous section with the maximum likelihood estimators under ranked set sampling (RSS) by Stokes [45]. Recall that the maximum likelihood estimators of μ and σ from RSS are respectively represented by $\hat{\mu}^*_{ML}$ and $\hat{\sigma}^*_{ML}$. Define the precision of the proposed unbiased estimators of μ under MRSS relative to the RSS maximum likelihood estimators of μ by

$$RP(\hat{\mu}_{UB_i}, \hat{\mu}^*_{ML}) = \frac{1}{I^*_{mn}(\mu) Var(\hat{\mu}_{UB_i})},$$

where $I^*_{mn}(\mu)$ is the Fisher information about μ in the RSS and i = 1 or 2 according as odd set size or even set size is being considered. From Stokes [45], we have in the case of the normal distribution

$$I^{*}_{mn}(\mu)=\frac{mn}{\sigma^2}+\frac{mn(n-1)}{\sigma^2}(0.4805)$$

Thus for odd and even set sizes, the respective relative precision relations are

$$RP(\hat{\mu}_{UB_1},\hat{\mu}^{*}_{ML})=\frac{1}{[1+(n-1)(0.4805)]v_{(p:n)}}$$

and

$$RP(\hat{\mu}_{UB_2},\hat{\mu}^{*}_{ML})=\frac{2}{[1+(n-1)(0.4805)][v_{(q:n)}+v_{(q+1:n)}]}$$

Similarly, define

$$RP(\hat{\sigma}_{UB_1},\hat{\sigma}^{*}_{M})=\frac{1}{I^{*}_{mn}(\sigma)Var(\hat{\sigma}_{UB_1})}$$

For the normal distribution, we have for even set sizes,

$$RP(\hat{\sigma}_{UB_2},\hat{\sigma}^{*}_{ML})=\frac{2[\alpha_{(q:n)}\alpha_{(q+1:n)}]^2}{[2+(n-1)(0.2705)][\alpha^2_{(q:n)}v_{(q+1:n)}+\alpha^2_{(q+1:n)}v_{(q:n)}]}$$

Under the exponential distribution, we have

$$RP(\hat{\sigma}_{UB_2},\hat{\sigma}^{*}_{ML})=\frac{[\alpha_{(p:n)}]^2}{[1+(n-1)(0.4041)]v_{(p:n)}}$$

and

$$RP(\hat{\sigma}_{UB_2},\hat{\sigma}^{*}_{ML})=\frac{2[\alpha_{(q:n)}\alpha_{(q+1:n)}]^2}{[1+(n-1)(0.4041)][\alpha^2_{(q:n)}v_{(q+1:n)}+\alpha^2_{(q+1:n)}v_{(q:n)}]}$$

for even and odd set sizes respectively.

We find from equation (3.35) that the Fisher information for σ from RSS under the gamma (2) and gamma (3) distributions are respectively given by

$$I^{*}_{mn}(\sigma)=\frac{2mn}{\sigma^2}+\frac{mn(n-1)}{\sigma^2}(0.87504)$$

and

$$I^*_{mn}(\sigma) = \frac{3mn}{\sigma^2} + \frac{mn(n-1)}{\sigma^2}(1.357).$$

The corresponding relative precision can therefore be found as in the cases considered above. We tabulate below, the results for all the distributions considered.

TABLE 4.4. The RELATIVE PRECISION of the PROPOSED UNBIASED ESTIMATORS with RESPECT to the RSS MAXIMUM LIKELIHOOD ESTIMATORS

Distribution	Relative Precision	Set size 2	3	4	5	6	7	8
N (μ, 1)	$RP(\hat{\mu}_{UB}, \hat{\mu}^*_{ML})$	0.99084	1.13657	1.13628	1.19315	1.19370	1.22372	1.22428
N (0,σ^2)	$RP(\hat{\sigma}_{UB}, \hat{\sigma}^*_{ML})$	0.20566		0.08722		0.04921		0.03191
exp (σ)	$RP(\hat{\sigma}_{UB}, \hat{\sigma}^*_{ML})$	0.91569	1.06353	1.03775	1.09790	1.08538	1.11785	1.11046
Gamma(2)	$RP(\hat{\sigma}_{UB}, \hat{\sigma}^*_{ML})$	0.95391	1.10152	1.08896	1.14844	1.14270	1.17461	1.17131
Gamma(3)	$RP(\hat{\sigma}_{UB}, \hat{\sigma}^*_{ML})$	1.08540	1.33736	1.38657	1.50596	1.53601	1.60567	1.62567

We observe from the table that beside the unbiased estimator for σ, all the other unbiased estimators, do better than the maximum likelihood estimator under RSS for set sizes greater than two, improving with increasing set size.

CHAPTER 5

MAXIMUM LIKELIHOOD ESTIMATION UNDER EXTREME RANKED SET SAMPLING

In this chapter, we consider maximum likelihood estimation under one modification of RSS namely extreme ranked set sampling (ERSS) as described in chapter 2. The motivation for this part of the work is the poor performance of MRSS in estimating the normal variance. All work here is done based on the assumption that the usual regularity conditions hold.

Of interest in this work are the distributions of the smallest and largest order statistics as well as that of the median from an odd set size. The respective density functions of the smallest and largest order statistics are

$$f_1(x) = n[1-F(x)]^{n-1} f(x) \tag{5.1}$$

and

$$f_n(x) = nF^{n-1}(x)f(x) \tag{5.2}$$

The density function of the median from an odd sized sample is

$$f_p(x) = \frac{1}{B\left(\frac{n+1}{2}, \frac{n+1}{2}\right)} F^{(n-1)/2}(x)[1-F(x)]^{(n-1)/2} f(x) \tag{5.3}$$

We now estimate the parameters of the location-scale family of distributions using the method of maximum likelihood estimation under ERSS following the methods of chapter 3.

5.1. MLE with Even Set Size

Suppose that we have an even set size and the m-cycle ERSS

$$\{X_{i(1)j}; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\} \bigcup \{X_{i(n)j}; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\}$$

from a population with cumulative distribution function $F\left(\frac{x-\mu}{\sigma}\right)$ and probability density function $\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$, where $X_{i(.)j}$ is the $(.)^{th}$ smallest or largest observation from the i^{th} set of j^{th} cycle. Then the loglikelihood function of the ERSS is

$$L_{E_1} = \ln \prod_{j=1}^{m} \left\{ \prod_{i=1}^{n/2} f_1(Z_{i(1)j}) \prod_{i=(n/2)+1}^{n} f_n(Z_{i(n)j}) \right\}, \tag{5.4}$$

where $Z_{i(.)j} = \dfrac{X_{i(.)j} - \mu}{\sigma}$.

$$L_{E_1} = K - mn \ln \sigma + \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n/2} \left[(n-1) \ln[1 - F(Z_{i(1)j})] + \ln f(Z_{i(1)j}) \right] \right.$$

$$\left. + \sum_{i=(n/2)+1}^{n} \left[(n-1) \ln F(Z_{i(n)j}) + \ln f(Z_{i(n)j}) \right] \right\}. \tag{5.5}$$

Suppose that σ is known. Then the maximum likelihood estimator of μ, $\hat{\mu}_{ML_{E_1}}$ is the solution of the equation

$$\sum_{j=1}^{m}\left\{\sum_{i=1}^{n/2}\left[(n-1)\frac{f(Z_{i(1)j})}{1-F(Z_{i(1)j})}-\frac{f'(Z_{i(1)j})}{f(Z_{i(1)j})}\right]-\sum_{i=(n/2)+1}^{n}\left[(n-1)\frac{f(Z_{i(n)j})}{F(Z_{i(n)j})}+\frac{f'(Z_{i(n)j})}{f(Z_{i(n)j})}\right]\right\}=0. \tag{5.6}$$

That is by differentiating equation (5.5) with respect to μ and setting it equal to zero. Upon finding the second derivative of (5.5) with respect to μ and taking the negative expectation, we obtain the Fisher information for μ from the ERSS as

$$\begin{aligned} I_{mn}^{E_1}(\mu)=&\frac{mn}{2\sigma^2}E\left\{\left[\left(\frac{f'(z)}{f(z)}\right)^2-\frac{f''(z)}{f(z)}\right][h_1(z)+h_2(z)]\right\}\\ &+\frac{mn(n-1)}{2\sigma^2}E\left\{\left[\frac{f'(z)}{1-F(z)}+\left(\frac{f(z)}{1-F(z)}\right)^2\right]h_1(z)-\left[\frac{f'(z)}{F(z)}-\left(\frac{f(z)}{F(z)}\right)^2\right]h_2(z)\right\}\end{aligned} \tag{5.7}$$

where $h_1(z)=n[1-F(z)]^{n-1}$ and $h_2(z)=nF^{(n-1)}(z)$. The Fisher information for μ from a SRS of size mn (see Stokes [49]) is

$$I_{mn}(\mu)=\frac{mn}{\sigma^2}E\left\{\frac{f'(z)}{f(z)}\right\}^2 \tag{5.8}$$

provided $E\left[\frac{\partial}{\partial\mu}\ln f(z)\right]=0$. Thus we can find the asymptotic relative precision of the ERSS estimator relative to that of the SRS, $\hat{\mu}_{ML}$, by

$$\lim_{m\to\infty} RP(\hat{\mu}_{ML_{E_1}},\hat{\mu}_{ML})=\frac{I_{mn}^{E_1}(\mu)}{I_{mn}(\mu)} \tag{5.9}$$

Similarly, if μ is known, then the ERSS maximum likelihood estimator of σ, $\hat{\sigma}_{ML_{E_1}}$ is the solution of the equation

$$mn-\sum_{j=1}^{m}\left\{\sum_{i=1}^{n/2}\left[(n-1)\frac{Z_{i(1)j}f(Z_{i(1)j})}{1-F(Z_{i(1)j})}-\frac{Z_{i(1)j}f'(Z_{i(1)j})}{f(Z_{i(1)j})}\right]\right.$$
$$\left.-\sum_{i=(n/2)+1}^{n}\left[(n-1)\frac{Z_{i(1)j}f(Z_{i(n)j})}{F(Z_{i(n)j})}+\frac{Z_{i(1)j}f'(Z_{i(n)j})}{f(Z_{i(n)j})}\right]\right\}=0 \tag{5.10}$$

and the Fisher information for σ from the ERSS is

$$I_{mn}^{E_1}(\sigma)=\frac{mn}{2\sigma^2}E\left\{\left[\left(\frac{zf'(z)}{f(z)}\right)^2-\frac{z^2f''(z)+2zf'(z)}{f(z)}\right][h_1(z)+h_2(z)]-2\right\}$$
$$+\frac{mn}{2\sigma^2}E\left\{(n-1)\left[\frac{z^2f'(z)+2zf(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right]h_1(z)\right.$$
$$\left.-(n-1)\left[\frac{z^2f'(z)+2zf(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right]h_2(z)\right\}, \tag{5.11}$$

where h_1 and h_2 are as previously defined. The Fisher information for σ from a SRS of size mn is

$$I_{mn}(\sigma)=\frac{mn}{\sigma^2}E\left\{\left[\frac{zf'(z)}{f(z)}\right]^2-1\right\} \tag{5.12}$$

Hence, we find the asymptotic relative precision of the ERSS estimator to that of the SRS by

$$\lim_{m\to\infty} RP(\hat{\sigma}_{ML_{E_1}},\hat{\sigma}_{ML})=\frac{I_{mn}^{E_1}(\sigma)}{I_{mn}(\sigma)} \tag{5.13}$$

5.2. MLE with Odd Set Size

Suppose now that we have an odd set sized ERSS with m cycles from the location-scale family of distributions. Then the ERSS is the set

$$\left\{X_{i(1)j}; i=1,2,\ldots,\frac{n-1}{2}; j=1,2,\ldots,m\right\} \bigcup \left\{X_{i(n)j}; i=\frac{n-1}{2}+1,2,\ldots,n-1; j=1,2,\ldots,m\right\} \bigcup \left\{X_{n(\frac{n+1}{2})j}; j=1,2,\ldots,m\right\}.$$

The log likelihood function of the ERSS is

$$L_{E_2} = K' - mn\ln\sigma + \sum_{j=1}^{m}\left\{\sum_{i=1}^{n'}\left[(n-1)\ln[1-F(Z_{i(1)j})] + \ln f(Z_{i(1)j})\right] + \sum_{i=n'+1}^{n-1}\left[(n-1)\ln F(Z_{i(n)j}) + \ln f(Z_{i(n)j})\right]\right\}$$
$$+\sum_{j=1}^{m}\left\{\left(\frac{n-1}{2}\right)\left[\ln[1-F(Z_{n(p)j})] + \ln F(Z_{n(p)j})\right] + \ln f(Z_{n(p)j})\right\}, \tag{5.14}$$

where $n' = (n-1)/2$ and $p = (n+1)/2$.

Suppose that σ is known. Then by differentiating equation (5.14) with respect to μ and setting the result it equal to zero, we obtain the maximum likelihood estimator of μ as the solution of the equation

$$\sum_{j=1}^{m}\left\{\sum_{i=1}^{n'}\left[(n-1)\frac{f(Z_{i(1)j})}{1-F(Z_{i(1)j})} - \frac{f'(Z_{i(1)j})}{f(Z_{i(1)j})}\right] - \sum_{i=n'+1}^{n-1}\left[(n-1)\frac{f(Z_{i(n)j})}{F(Z_{i(n)j})} - \frac{f'(Z_{i(n)j})}{f(Z_{i(n)j})}\right]\right\}$$
$$+\sum_{j=1}^{m}\left\{\left(\frac{n-1}{2}\right)\left[\frac{f(Z_{ni(1)j})}{1-F(Z_{n(1)j})} - \frac{f(Z_{n(1)j})}{F(Z_{n(1)j})}\right] - \frac{f'(Z_{n(1)j})}{f(Z_{n(1)j})}\right\} = 0 \tag{5.15}$$

The Fisher information for μ from the ERSS is

$$I_{mn}^{E_2}(\mu) = \frac{mn}{\sigma^2}E\left\{\frac{1}{n}\left[\frac{n-1}{2}[h_1(z)+h_2(z)] + g(z)\right]\left[\left(\frac{f'(z)}{f(z)}\right)^2 - \frac{f''(z)}{f(z)}\right]\right\}$$
$$+\frac{mn(n-1)}{2\sigma^2}E\left\{\frac{1}{n}((n-1)h_1(z)+g(z))\left[\frac{f'(z)}{1-F(z)} + \left(\frac{f(z)}{1-F(z)}\right)^2\right]\right.$$
$$\left. - \frac{1}{n}((n-1)h_2(z)+g(z))\left[\frac{f'(z)}{F(z)} + \left(\frac{f(z)}{F(z)}\right)^2\right]\right\} \tag{5.16}$$

where $h_1(z)$ and $h_2(z)$ are as previously defined and

$$g(z) = \frac{1}{B(p,n-p+1)}F^{p-1}(z)[1-F(z)]^{n-p}$$

with $p=(n+1)/2$.

Therefore, the asymptotic relative precision of the ERSS estimator with respect to that of the SRS is

$$\lim_{m\to\infty} RP(\hat{\mu}_{ML_{E_2}},\hat{\mu}_{ML})=\frac{I_{mn}^{E_2}(\mu)}{I_{mn}(\mu)} \tag{5.17}$$

Similarly, the ERSS maximum likelihood estimator of σ, $\hat{\sigma}_{ML_{E_2}}$, is the solution of the equation

$$-mn+\sum_{j=1}^{m}\left\{\sum_{i=1}^{n'}\left[(n-1)\frac{Z_{i(1)j}f(Z_{i(1)j})}{1-F(Z_{i(1)j})}-\frac{Z_{i(1)j}f'(Z_{i(1)j})}{f(Z_{i(1)j})}\right]-\sum_{i=n'+1}^{n-1}\left[(n-1)\frac{Z_{i(n)j}f(Z_{i(n)j})}{F(Z_{i(n)j})}-\frac{Z_{i(1)j}f'(Z_{i(n)j})}{f(Z_{i(n)j})}\right]\right\}$$
$$+\sum_{j=1}^{m}\left\{\left(\frac{n-1}{2}\right)\left[\frac{Z_{n(p)j}f(Z_{n(p)j})}{1-F(Z_{n(p)j})}-\frac{Z_{n(p)j}f(Z_{n(p)j})}{F(Z_{n(p)j})}\right]-\frac{Z_{n(p)j}f'(Z_{n(p)j})}{f(Z_{n(p)j})}\right\}=0$$

and the ERSS Fisher information for σ is

$$I_{mn}^{E_2}(\sigma)=\frac{mn}{\sigma^2}E\left\{\frac{1}{n}\left[\left(\frac{zf'(z)}{f(z)}\right)^2-\frac{z^2f''(z)+2zf'(z)}{f(z)}\right]\left((\frac{n-1}{2})[h_1(z)+h_2(z)]+g(z)\right)-1\right\}$$
$$+\frac{mn(n-1)}{2\sigma^2}E\left\{\frac{1}{n}\left[\frac{z^2f'(z)+2zf(z)}{1-F(z)}+\left(\frac{zf(z)}{1-F(z)}\right)^2\right][(n-1)h_1(z)+g(z)]\right. \tag{5.18}$$
$$\left.-\frac{1}{n}\left[\frac{z^2f'(z)+2zf(z)}{F(z)}+\left(\frac{zf(z)}{F(z)}\right)^2\right][(n-1)h_2(z)+g(z)]\right\}$$

Hence the asymptotic relative precision is

$$\lim_{m\to\infty} RP(\hat{\sigma}_{ML_{E_2}},\hat{\sigma}_{ML})=\frac{I_{mn}^{E_2}(\sigma)}{I_{mn}(\sigma)} \tag{5.19}$$

5.3. Comparisons with the RSS and MRSS Estimators

In this section we present a comparison of the ERSS maximum likelihood estimators with their RSS and MRSS counterparts obtained in Chapter 3 and also with some of the non-parametric estimators. We limit our comparisons to the cases of the estimation of the normal mean and variance, and then the exponential mean.

Table 5.1 shows the asymptotic relative precision of the maximum likelihood estimators of the normal mean under each of ERSS, MRSS obtained in Chapter 3 and RSS (Stokes [45]). It also shows the relative precision values for the non-parametric RSS (Dell and Clutter [10]), MRSS (Muttlak [25]) and ERSS estimators computed from Samawi et al. [36]. We observe here that even though the ERSS maximum likelihood estimator is an improvement of its non-parametric counterpart, it does not do as well as the parametric and non-parametric RSS and MRSS estimators.

TABLE 5.1. A COMPARISON of ERSS MAXIMUM LIKELIHOOD ESTIMATORS of the NORMAL MEAN with the other ESTIMATORS

set size	Asymptotic relative precision (Maximum Likelihood)			Relative precision Non-Parametric methods		
n	RSS	MRSS	ERSS	RSS	MRSS	ERSS
2	1.48	1.48	1.48	1.47	1.47	1.47
3	1.96	2.23	1.96	1.91	2.23	1.91
4	2.44	2.78	2.10	2.35	2.77	2.03
5	2.92	3.49	2.56	2.77	3.49	2.41
6	3.40	4.07	2.53	3.19	4.06	2.40
7	3.88	4.75	3.00	3.59	4.75	2.73
8	4.36	5.34	2.87	4.00	5.34	2.68

In Table 5.2, we present the asymptotic relative precision for the maximum likelihood estimators of the normal variance under each of RSS (Stokes [45]), MRSS as obtained in Chapter 3 and ERSS. The asymptotic relative precision for the non-parametric estimator from Stokes [42] is also shown in the table. The ERSS maximum likelihood estimator is clearly seen to dominate all the other estimators except for a set size of 2, when together with the MRSS estimator, it coincides with the RSS maximum likelihood estimator.

TABLE 5.2. A COMPARISON of ESTIMATORS of the σ FROM a NORMAL DISTRIBUTION

set size	Asymptotic relative precision (Maximum Likelihood)			Non-Parametric methods
n	RSS	MRSS	ERSS	RSS
2	1.14	1.14	1.14	1.00
3	1.27	0.98	1.27	1.08
4	1.41	1.08	1.74	1.18
5	1.54	0.98	1.85	1.27
6	1.68	1.05	2.40	1.38
7	1.81	0.99	2.49	1.48
8	1.95	1.04	3.06	1.57

Table 5.3 compares the maximum likelihood and non-parametric estimators of the exponential mean under the three schemes.

TABLE 5.3. A COMPARISON OF ESTIMATORS OF THE EXPONENTIAL MEAN

set size	Asymptotic relative precision (Maximum Likelihood)			Relative precision Non-Parametric methods		
n	RSS	MRSS	ERSS	RSS	MRSS	ERSS
2	1.40	1.40	1.40	1.33	1.33	1.33
3	1.81	1.92	1.81	1.64	2.25	1.64
4	2.21	2.37	2.06	1.92	2.44	1.17
5	2.62	2.88	2.44	2.19	2.23	1.32
6	3.02	3.33	2.58	2.45	2.14	0.75
7	3.42	3.83	2.96	2.70	1.80	0.81
8	3.83	4.29	3.03	2.94	1.67	0.46

The relative precision of the RSS, MRSS and ERSS were respectively obtained from Dell and Clutter [10], Muttlak [25] and Samawi et al. [36]. The ERSS maximum likelihood estimator is clearly an improvement of its non-parametric counterpart. It also dominates the non-parametric RSS estimator. Just as in the case of the MRSS maximum likelihood estimator, the ERSS maximum likelihood estimator dominates the non-parametric MRSS estimator except for set sizes of 2 and 3. The ERSS maximum likelihood estimator does not however do as well as the RSS and MRSS estimators.

CHAPTER 6

SUMMARY and DISCUSSION

In chapter 3, we considered the estimation of the location and scale parameters using the method of maximum likelihood under median ranked set sampling (MRSS). We used various distributions to illustrate our method and drew comparisons with the results obtained using maximum likelihood estimation under ranked set sampling (RSS) (Stokes [45]), non-parametric ranked set sampling (Dell and Clutter [10]), and non-parametric MRSS (Muttlak [25]).

In the estimation of all the parameters beside the normal variance, the maximum likelihood estimator under MRSS is seen to do better than that under RSS and that under the non-parametric RSS. The maximum likelihood estimator of the normal mean under MRSS is exactly as efficient as that under the non-parametric MRSS. The asymptotic relative precision values obtained coincide with the relative precision values obtained in the non-parametric case, then the non-parametric MRSS estimator of the normal mean also has the lower bound variance. Thus, the non-parametric MRSS estimator of the normal mean is a best linear unbiased estimator (BLUE).

For the normal variance, the MRSS maximum likelihood estimator was found to be worse than its simple random sample counterpart for odd set sizes. For even set sizes, it does its best when it coincides with RSS at a set size of 2. This may indicate that the MRSS will in general not be a successful method with regard to the estimation of the normal variance. Sinha et al. [40] did point out this in using the traditional median, which coincides with the MRSS for odd set sizes.

In estimating the scale parameter of the exponential and gamma distributions, we observed that the MRSS maximum likelihood estimator is dominant over all other methods except at set sizes of 3 and 4.

We proposed some unbiased estimators of the location and scale parameters under MRSS in Chapter 4, and compared them with the corresponding maximum likelihood estimators. The results here revealed that these estimators under the distributions considered are nearly as efficient as the maximum likelihood estimators are, and can thus be used for computational convenience. The price to pay however is the slight loss in efficiency. One advantage of the proposed unbiased estimators of Chapter 4 over the maximum likelihood estimators of chapter 3 is that the results of the maximum likelihood estimators are asymptotic (i.e. the relative precision values hold for large cycle size, m) whereas those of the proposed unbiased estimators are not.

Comparing the proposed unbiased estimators of Chapter 4 to the RSS maximum likelihood estimators, they are found to be doing better in the case of the normal mean for set sizes greater than 2. The case of the normal variance is however very bad for the proposed unbiased estimators. The proposed estimators were also shown to dominate the

RSS maximum likelihood estimators in the estimation of the scale parameters of the exponential and gamma distributions for set sizes greater than 2.

The woeful performance of the MRSS maximum likelihood estimator of σ for the normal distribution motivated the consideration of maximum likelihood estimation under ERSS in Chapter 5. Results here show the ERSS to be a success story in the estimation of the normal standard deviation. It indeed dominates all the other schemes considered. However, it does not do as well as the other schemes in all the other cases considered.

In conclusion, the work of this thesis shows that in estimating the normal mean, μ, the best estimators among the estimators considered in this are the MRSS non-parametric and maximum likelihood estimators. Since they yield the same relative precision, it is highly recommended to use the non-parametric estimators for computational convenience. On the other hand the best estimator for the normal standard deviation is the ERSS maximum likelihood estimator. In case of estimating the scale parameters of the exponential and gamma distributions, the MRSS maximum likelihood estimators are the best candidates.

Finally, we wish to recommend the application of other modifications of RSS in the direction of the work of this thesis. For instance, it will be useful to consider the partial schemes of RSS, MRSS and ERSS. Success in such projects will indicate that we do not need to take as large a sample as we did in this thesis to achieve comparable or even better efficiencies. The non-parametric estimation of the normal variance using ERSS will also be an interesting problem to consider. Also, a consideration of general methods to simplify the estimators of chapter 3 will be worth while. Finally, a generalization of the methods of this thesis to all kinds of distributions rather the location-scale family alone will be useful.

REFERENCE

1. Abu-Dayeh, W. and Muttlak, H. A. (1996). Using ranked set sampling for hypothesis tests on the scale parameter of the exponential and uniform distributions. *Pakistan Journal of Statistics*, **12**(2), 131-38.

2. Bohn, Lora L. (1996). A review of non-parametric ranked set sampling methodology. *Communication in Statistics-Theory and Methods*. **25**(11), 2675-85.

3. Bohn, Lora L. and Wolfe, D. A. (1992). Non-parametric two-sample procedures for ranked set samples data. *Journal of the American Statistical Association*, **87**(418), 552-561.

4. Bohn, Lora L. and Wolfe, D. A. (1994). The effect of imperfect judgement rankings on properties of procedures based on the ranked set samples analog of the Mann-Whitney-Wilcoxon statistic. *Journal of the American Statistical Association*, **89**(425), 168-76.

5. Bhoj, D. S. (1997). Estimation of parameters of the extreme value distribution using ranked set sampling. *Communications in Statistic-Theory methods*, **26**(3), 653-67.

6. Bhoj, D. S. and Ahsanullah, M. (1996). Estimation of parameters of the Generalized Geometric Distribution using ranked set sampling. *Biometrics*, **52**, 685-94.

7. Chen. S. –H. (1983). Ranked set sampling theory with selective probability vector. *Journal of Statistical Planning and Inference*, **8**, 161-174.

8. Cochran, W. G. (1977). Sampling techniques. Wiley series in probability and mathematical statistics

9. David, H. A. (1981). Order statistics (Second Edition). Wiley series in probability and statistics.

10. Dell, D. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics bacground. *Biometrics*, **28**, 545-55.

11. Halls, L. K. and Dell, T. R. (1966). Trial of ranked set sampling for forage yields. *Forest Science*, **12**(1).

12. Harter, H. L. and Balakrishnan, N. (1996). CRC handbook of tables for the use of order statistics in estimation, *CRC Press*, Boca Raton.

13. Hossain, S. S. and Muttlak, H. A. (1999). Paired ranked set sampling: A more efficient procedure. *Environmetrics*, **10**, 195-212.

14. Koti, K. A. and Babu, G. J. (1996). Sign test for ranked set sampling. *Communications in Statistics-Theory and Methods*. 25(7), 1617-30.

15. Kvam, P. H. and Samaniego, F. J. (1993). On the inadmissibility of empirical averages as estimators in ranked set sampling. *Journal of Statistical Planning and Inference*, 36, 39-55.

16. Kvam, P. H. and Samaniego, F. J. (1994). Non-parametric maximum likelihood estimation based on ranked set sample. *Journal of the American Statistical Association*, 89(426), 526-37.

17. Lam, K, Sinha, B. K and Wu, Z (1994). Estimation of Parameters in a Two-parameter Exponential Distribution using Ranked Set Sample. *Annals of the Institute of Statistical Mathematics*, 46(4), 723-36.

18. Lloyd, E. H. (1952). Least-squares estimation of location and scale parameters using order statistics. *Biometrika* 39, 88-95.

19. McIntyre, G. A. (1952). A method of unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 3, 385-390

20. Muttlak, H. A. (1995). Parameters estimation in a simple linear regression using rank set sampling. *The Biometrical Journal*, 37(7), 799-810.

21. Muttlak, H. A. (1995). Combining the line intercept sampling with the ranked set sampling. *Journal of Information & Optimization Science*, 16(1), 1-16.

22. Muttlak, H. A. (1996). Estimation of parameters for one-way layout with rank set sampling. *The Biometrical Journal*, 38(4), 507-515

23. Muttlak, H. A. (1996). Pair rank set sampling. *The Biometrical Journal*, 38(7), 507-515.

24. Muttlak, H. A. (1996). Estimation of parameters in a multiple regression model using rank set sampling. *Journal of Information & Optimization Science*, 17(3), pp1-16.

25. Muttlak, H. A. (1997). Median ranked set sampling. *Journal of Applied Statistical Science*, 6(4), 245-55.

26. Muttlak, H. A. (1998). Median ranked set sampling with size biased probability of selection. *The Biometrical Journal*, 40(4), 455-65.

27. Muttlak, H. A. (1998). Median ranked set sampling with concomitant variables and comparison with rank set sampling and regression estimators. *Environmetrics*, 9, 255-67.

28. Muttlak, H. A. and Abu-Dayyeh, W. (1998). Testing some hypotheses about the normal distribution using ranked set sample: a more powerful test. *Journal of Information & Optimization Science*, **19**(1), 1-11.

29. Muttlak, H. A. and McDonald, L. L. (1990). Ranked set sampling with size-biased probability of selection. *Biometrics*, **46**, 435-45.

30. Muttlak, H. A. and McDonald, L. L. (1992). Ranked set sampling and the line intercept method: a more efficient procedure. *The Biometrical Journal*, **46**, 435-45.

31. Patil, G. P., Sinha, A. K. and Taillie, C. (1993). Relative precision of ranked set sampling: a comparison with the regression estimator. *Environmetrics*, 4(4), pp399-412.

32. Patil, G. P., Sinha, A. K. and Taillie, C. (1993). Ranked set sampling from a finite population in the presence of a trend on site. *Journal of Applied Statistical Science*, 1(1), 51-65.

33. Patil, G. P., Sinha, A. K. and Taillie, C. (1994). Ranked set sampling. *A Handbook of Statistics*, vol. 12, pp167-200.

34. Patil, G. P., Sinha, A. K. and Taillie, C. (1997). Ranked set sampling, coherent rankings and size-biased permutations. *Journal of Statistical Planning and Inferecce*, **63**, 311-324.

35. Ridout, M. S. and Cobby, J. M., (1987). Ranked set sampling with non-random selection of sets and errors in ranking. *Applied Statistics*, 36(2), 145-52.

36. Samawi, H. M., Ahmed, M. S. and Abu-Dayyeh. Estimating the population mean using extreme raked set sampling. *The Biometrical Journal*, 38(5), 577-86.

37. Samawi, H. M. and Muttlak, H. A. (1996). Estimation of ratio using rank set sampling. *The Biometrical Journal*, 38(6), 753-64.

38. Shen, W. -H. (1994). Use of rank set sampling for test of normal mean. *Calcutta Statistical Association Bulletin*. **44**, Nos. 175-176.

39. Shirahata, S. (1982). An extension of the ranked set sampling theory. *Journal of Statistical Planning and Inference*, **6**, 65-72.

40. Sinha, B. K., B. K. Sinha, and Purkayastha, S. (1996), On Some Aspects of Ranked Set Sampling for Estimation of Normal and Exponential Parameters. *Statistics and Decisions*, **14**, 223-240.

41. Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics-Theory and Methods*, A6(12), 1207-11.

42. Stokes, S. L. (1980), Estimation of Variance Using Judgement Ordered Ranked Set Samples. *Biometrics*, **36**, 35-42.

43. Stokes, S. L. (1980). Inference on the correlation coefficient in bivariate normal populations from ranked set samples. *Journal of the American Statistical Association*, **75**(372).

44. Stokes, S. L. (1986). Ranked set sampling. *Encyclopedia of Statistical Sciences*, vol. 7, S. Kotz, N. L. Johnson and C. B. Read (eds), Wiley, New York, 585-88.

45. Stokes, S. L. (1995). Parametric Ranked Set Sampling. *Annals of the Institute of Statistical Mathematics*. **47**(3), 465-482.

46. Stokes, S. L. and Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimation distribution functions. *Journal of the American Statistical Association*, **83**(402).

47. Takahasi, K. (1969). On the estimation of the population mean based on ordered samples from eqricorrelated multivariate distribution. *Annals of the Institute of Statistical Mathematics*, **21**, 249-55.

48. Takahasi, K. (1970). Practical note on estimation of population means based on samples stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, **21**, 421-8.

49. Takahasi, K. and Futasuya, M. (1998). Dependence between order statistics in samples from finite population and its application to ranked set sampling. *Annals of the Institute of Statistical Mathematics*, **50**(1), 49-70.

50. Takahasi K. and Wakimoto K. (1968). On Unbiased Estimates Of The Population Mean Based On The Sample Stratified By Means Of Ordering *Annals of the Institute of Statistical Mathematics*, **21**, 249-55.

51. Yanagawa, T and Chen, S. –H. (1980). The MG-procedure in ranked set sampling. *Journal of Statistical Planning and Inference*, **4**, 33-44.

52. Yanagawa, T. and Shirahata, S. (1976). Ranked set sampling theory with selective probability matrix. *Australian Journal of Statistics*, **18**(1,2), 45-52.

53. Yu, Philip L. H. and Lam, K. (1997). Regression estimator in ranked set sampling. *Biometrics*, **53**, 1070-80.

VITA

In December 1994, I graduated from the University of Ghana with a Bachelor of Science Honors degree in Mathematics and Statistics and served in the Mathematics Department of the same university as a Teaching Assistant on National Service. In October 1995, I was appointed a Tutor of Mathematics at Herman Gmeiner International College, Tema, Ghana. I joined the Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals as a graduate student on Research Assistantship in April 1997. The successful defense of this thesis in December 1999 marks my acquisition of the Master of Science degree in Mathematics.